

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 15:08:33 ; Search time 694 Seconds

(without alignments)
3184.733 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047
Sequence: 1 MVKLHTLADHGDVNCACF.....LTPNRLKMAINRWLETHQK 384

Scoring table:

BLISSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 433806 segs, 2877871033 residues

Total number of hits satisfying chosen parameters: 3483234

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cg2_1/USPTO.spool/US10077111/runat_04022005_121142_7236/app.query.fasta_1.583
-DB=Published Applications NA -QFMT=fastlap -SUFFIX=p2nszlm200.rmpb
-MINMATCH=0.1 -LOOPEXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blssum62 -TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct
-THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext
-HEAPSIZ=500 -MINLEN=0 -MAXLEN=200
-USER=US10077111_@CG2_1_1_628 @runat_04022005_121142_7236 -NCPU=6 -ICPU=3
-NO MMP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA: *
1: /cg2_1/USPTO.spool/US10077111/runat_04022005_121142_7236/app.query.fasta_1.583
2: /cg2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
3: /cg2_6/prodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cg2_6/prodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cg2_6/prodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cg2_6/prodata/2/pubpna/US07_PUBCOMB.seq.*
7: /cg2_6/prodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cg2_6/prodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cg2_6/prodata/2/pubpna/US09_PUBCOMB.seq.*
10: /cg2_6/prodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cg2_6/prodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cg2_6/prodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cg2_6/prodata/2/pubpna/US10_PUBCOMB.seq.*
14: /cg2_6/prodata/2/pubpna/US10_PUBCOMB.seq.*
15: /cg2_6/prodata/2/pubpna/US10C_PUBCOMB.seq.*
16: /cg2_6/prodata/2/pubpna/US10D_PUBCOMB.seq.*
17: /cg2_6/prodata/2/pubpna/US10E_PUBCOMB.seq.*
18: /cg2_6/prodata/2/pubpna/US10F_PUBCOMB.seq.*
19: /cg2_6/prodata/2/pubpna/US10_NEW_PUB.seq.*
20: /cg2_6/prodata/2/pubpna/US11_NEW_PUB.seq.*
21: /cg2_6/prodata/2/pubpna/US60_NEW_PUB.seq.*
22: /cg2_6/prodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| C 1 | 300 | 14.7 | 180 | US-09-864-761-27687 | Sequence 27687, A |
| C 2 | 210 | 10.3 | 132 | US-09-864-761-30542 | Sequence 30542, A |
| C 3 | 150 | 7.3 | 92 | US-10-029-386-16571 | Sequence 16571, A |
| C 4 | 93.5 | 4.6 | 154 | US-10-029-386-16571 | Sequence 16571, A |
| C 5 | 87 | 4.3 | 150 | US-10-029-386-16571 | Sequence 16571, A |
| C 6 | 82 | 4.0 | 173 | US-09-864-761-29419 | Sequence 29419, A |
| C 7 | 79.5 | 3.9 | 159 | US-10-305-720-417 | Sequence 417, A |
| C 8 | 78.5 | 3.8 | 155 | US-10-029-386-16111 | Sequence 16111, A |
| C 9 | 72.5 | 3.5 | 197 | US-09-998-598-205 | Sequence 205, A |
| C 10 | 72.5 | 3.5 | 197 | US-09-998-598-205 | Sequence 205, A |
| C 11 | 72.5 | 3.5 | 197 | US-09-998-598-205 | Sequence 205, A |
| C 12 | 71.5 | 3.5 | 187 | US-10-425-115-24663 | Sequence 24663, A |
| C 13 | 67.5 | 3.3 | 159 | US-10-424-599-124934 | Sequence 124934, A |
| C 14 | 65.5 | 3.2 | 167 | US-10-029-386-16501 | Sequence 16501, A |
| C 15 | 65 | 3.2 | 131 | US-10-060-036-1810 | Sequence 1810, A |
| C 16 | 64 | 3.1 | 118 | US-09-864-761-23956 | Sequence 23956, A |
| C 17 | 64 | 3.1 | 142 | US-10-029-386-16384 | Sequence 16384, A |
| C 18 | 63.5 | 3.1 | 142 | US-10-029-386-16305 | Sequence 16305, A |
| C 19 | 62.5 | 3.1 | 156 | US-10-424-599-37936 | Sequence 37936, A |
| C 20 | 62.5 | 3.1 | 156 | US-09-814-353-5015 | Sequence 5015, A |
| C 21 | 62 | 3.0 | 133 | US-09-814-353-11307 | Sequence 11307, A |
| C 22 | 61 | 3.0 | 123 | US-10-077-111-19 | Sequence 19, A |
| C 23 | 61 | 3.0 | 181 | US-10-425-115-136942 | Sequence 136942, A |
| C 24 | 61 | 3.0 | 181 | US-09-920-100A-1407 | Sequence 1407, A |
| C 25 | 61 | 3.0 | 181 | US-10-033-528-1407 | Sequence 1407, A |
| C 26 | 61 | 3.0 | 181 | US-10-099-926-1407 | Sequence 1407, A |
| C 27 | 61 | 3.0 | 189 | US-10-029-386-15506 | Sequence 15506, A |
| C 28 | 60.5 | 3.0 | 196 | US-10-424-599-42401 | Sequence 42401, A |
| C 29 | 60 | 2.9 | 177 | US-09-867-701-6419 | Sequence 6419, A |
| C 30 | 59.5 | 2.9 | 189 | US-09-738-626-3303 | Sequence 3303, A |
| C 31 | 59.5 | 2.9 | 103 | US-09-864-761-21318 | Sequence 21318, A |
| C 32 | 59 | 2.9 | 189 | US-09-922-293-2560 | Sequence 2560, A |
| C 33 | 59 | 2.9 | 103 | US-09-969-373-651 | Sequence 651, A |
| C 34 | 59 | 2.9 | 115 | US-09-987-899-6165 | Sequence 6165, A |
| C 35 | 59 | 2.9 | 177 | US-10-029-386-21065 | Sequence 21065, A |
| C 36 | 59 | 2.9 | 133 | US-10-029-386-21433 | Sequence 21433, A |
| C 37 | 58.5 | 2.9 | 183 | US-10-425-115-39652 | Sequence 39652, A |
| C 38 | 58.5 | 2.9 | 183 | US-10-425-115-39652 | Sequence 39652, A |
| C 39 | 58 | 2.8 | 165 | US-10-437-963-28604 | Sequence 28604, A |
| C 40 | 58 | 2.8 | 165 | US-10-437-963-28604 | Sequence 28604, A |
| C 41 | 57.5 | 2.8 | 145 | US-10-437-963-28604 | Sequence 28604, A |
| C 42 | 57.5 | 2.8 | 145 | US-10-437-963-28604 | Sequence 28604, A |
| C 43 | 57 | 2.8 | 141 | US-10-085-783A-54429 | Sequence 54429, A |
| C 44 | 57 | 2.8 | 141 | US-10-282-122A-5300 | Sequence 5300, A |
| C 45 | 57 | 2.8 | 151 | US-09-864-761-30033 | Sequence 30033, A |

ALIGNMENTS

RESULT 1
US-09-864-761-27687/c
Sequence 27687, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Weisheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aomicca-X-1
CURRENT APPLICATION NUMBER: US/09/864, 761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366

```

; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27687
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006501.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EST HUMAN HIT: AL040518.1, EVALUE 1.00e-94
; OTHER INFORMATION: NT HIT: U43139.1, EVALUE 1.10e-01
; OTHER INFORMATION: SWISSPROT HIT: Q16760, EVALUE 6.80e-02
US-09-864-761-27687

Alignment Scores:
Pred. No.: 2,16e-26 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14,66% Indels: 0
DB: 9 Gaps: 0

US-10-077-111-13 (1-384) x US-09-864-761-27687 (1-180)

QY 227 ArgArgThnGluHnIsgInleuYsgInphethnGluAspTpsErgGluIuValIser 246
Db 178 AGGCGCACAGAACATCAGCTGAGACATTACCGAAGATTGTCAGAGGAGATGTCTCA 119
QY 247 ThTTPleuCYaAlaGlnAspleuLySAspleuValGlyIlePhelySmetAsnAnlle 266
Db 118 ACATGCGCTTGTGCACAAAGATTAAAGATTCTTGTGTAATTTTCAAGATGATTAACATT 59
QY 267 AAsGlyLySgIuLeuLeuAnleuThnLySgIuSerLeuAlaAspAspleuLySile 285
Db 58 GATGCAAAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAATTT 2

RESULT 2
```

```

US-09-864-761-30542/c
; Sequence 30542, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US 09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30542
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009307.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN HB1100, SIGNAL = 0.68
; OTHER INFORMATION: EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68
; OTHER INFORMATION: NT HIT: X71133.1, EVALUE 1.20e+00
; OTHER INFORMATION: SWISSPROT HIT: P54860, EVALUE 1.00e+00
US-09-864-761-30542

Alignment Scores:
Pred. No.: 7,04e-16 Length: 132
Score: 210.00 Matches: 43
```


Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 10.26%
DB: 9
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

US-10-077-111-13 (1-384) x US-09-864-761-30542 (1-132)

QY 290 LeuArgSerlyValLeuArglyValIleGluGluLeuArgThrlyValIleSerLeuSer 309
DB 130 CTGGGTGTAAGCGTGAAGGAAATTAAGAGCTCAGACCAAGTTAAATCCCTTCT 71

QY 310 SerGlyIleProArgGluPheIleCysProIleThrArgGluLeuMetlyAspProVal 329
DB 70 TCAGGAATTCCTGATGATTTATATGTCACATACTAGAGAACTTAAGAAAGATCCGCTC 11

QY 330 IleAlaSer 332
DB 10 ATCCCATCA 2

RESULT 3

US-10-106-698-3381
; Sequence 3381, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: P4005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; PRIOR FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 3381
; LENGTH: 92
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (45)..(45)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (78)..(78)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (86)..(86)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (89)..(89)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc_feature
; LOCATION: (91)..(91)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-3381

Alignment Scores:

Pred. No.: 5.64e-09 Length: 92
Score: 150.00 Matches: 26
Percent Similarity: 92.86% Conservative: 0
Best Local Similarity: 92.86% Mismatches: 2
Query Match: 7.33% Indels: 0
DB: 15 Gaps: 0

US-10-077-111-13 (1-384) x US-10-106-698-3381 (1-92)

QY 161 GlyAspLeuThrValITTPAspAPLyMeTarGySleuHisSerGlyULyAlaHisasp 180
DB 1 GGTGATTTAAACAAGTGTGGATGATTAATGAATGAGTGTCTGTCATAGNAAAAAGACATGAT 60

QY 181 LeuGlyIleThrCysCysAspPhe 188
DB 61 CTTGGAATTAACCTGCTGNGATTT 84

RESULT 4

US-10-029-386-26571
; Sequence 26571, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 26571
; LENGTH: 174
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR3.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.4
; OTHER INFORMATION: SWISSPROT HIT: O60907, EVALUE 3.00e-29
; OTHER INFORMATION: NT HIT: g14764641, EVALUE 7.00e-90
; OTHER INFORMATION: EST_HUMAN HIT: BE465313.1, EVALUE 1.00e-89
US-10-029-386-26571

Alignment Scores:

Pred. No.: 0.0841 Length: 174
Score: 93.50 Matches: 23
Percent Similarity: 49.12% Conservative: 5
Best Local Similarity: 40.35% Mismatches: 28
Query Match: 4.57% Indels: 1
DB: 16 Gaps: 1

US-10-077-111-13 (1-384) x US-10-029-386-26571 (1-174)

QY 71 SerCysSerThrAspGlyThrThrValIleThrAsnThrGluAsnGlyGlnMetLeuAla 90

DB 3 AGTCTCTGTTGATTTTCTACGTCGCGACTGTGGACCTAAGACGCGCTGCAACCAC 62

QY 91 ValMetGluGlnProSerGlySerProValArgValCysGlnPheSerProAspSerThr 110

DB 63 ACGCTCAGAGAGCATAG---GAGCCTGTCTATAGCGTAGCTTCAAGCCGATGGAG 119

QY 111 CysLeuAlaSerGlyAlaAlaAspGlyThrValValIleThrPheAlaGln 127

DB 120 TACTTGGCAGTGGATCTTCGACAAGTGGCTCATCTTCGAAATCTCAG 170

RESULT 5

US-10-029-386-18677
; Sequence 18677, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 18677
; LENGTH: 150
; TYPE: DNA

```

; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC022073.13
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: SWISSPROT HIT: Q92122, EVALUE 7.00e-26
; OTHER INFORMATION: NT HIT: g46005931, EVALUE 3.00e-79
; OTHER INFORMATION: EST_HUMAN HIT: BB892183.1, EVALUE 5.00e-79
US-10-029-386-18677

Alignment Scores:
Pred. No.: 0.396      Length: 150
Score: 87.00      Matches: 18
Percent Similarity: 56.76%      Conservative: 3
Best Local Similarity: 48.65%      Mismatches: 16
Query Match: 4.25%      Indels: 0
DB: 16      Gaps: 0

US-10-077-111-13 (1-384) x US-10-029-386-18677 (1-150)

QY 51 LvPheHieThrYrAlaValHAcGyGcYcPheSerProSerGIYHileLeuAla 70
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 9 AACGACACTTGTGCTCTATTCATCTGTGAGATTAGCTCTATGAGAACTCTATGCC 68

QY 71 SerCySerThrAaPgiYThrThrValLeuTrpAsnThrGIuAengIyGln 87
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 AGCGGTTCAGAAAGATGGAACATTGAGACTGTGCAACACTGTGATGAGAAAA 119

RESULT 6
US-09-864-761-29419
; Sequence 29419, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 29419
; LENGTH: 173
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC021976.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: SWISSPROT HIT: P42841, EVALUE 4.00e-08
; OTHER INFORMATION: NT HIT: AF069313.2, EVALUE 5.00e-93
; OTHER INFORMATION: EST_HUMAN HIT: H56093.1, EVALUE 8.00e-93
US-09-864-761-29419

Alignment Scores:
Pred. No.: 1.96      Length: 173
Score: 82.00      Matches: 18
Percent Similarity: 55.56%      Conservative: 12
Best Local Similarity: 33.33%      Mismatches: 22
Query Match: 4.01%      Indels: 2
DB: 9      Gaps: 1

US-10-077-111-13 (1-384) x US-09-864-761-29419 (1-173)

QY 36 LeuTYrSerLeuAArgAaPheThrGIuLeuPProHISerProLeuIyPheHieThrTYr 55
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 CTTGGAAATATGATTAATATACACACATGATA-----CGGAACCTAGAAAGACATCACCAT 60

QY 56 AlaValHieCyGcYcPheSerProSerGIYHileLeuAlaSerCySerThrAaP 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GATGTGTAGCTTGACCTTCTCTCTGATGAGACATTACGCTACCTCATCTTATGAT 120

QY 76 GIYThrThrValLeuTrpAsnThrGIuAengIyGlnMetLeu 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 ACTCGAGTATATATCTGGATCCACATTAATGAGACATTCTG 162

RESULT 7
US-10-305-720-417
; Sequence 417, Application US/10305720
; Publication No. US20040010136A1
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice K.; Seilhamer, Jeffrey J.
; TITLE OF INVENTION: Composition for the Detection of Signaling Pathway Gene Expression
; FILE REFERENCE: PA-0002-1 CON
; CURRENT APPLICATION NUMBER: US/10/305,720
; PRIOR FILING DATE: 2002-11-26
; PRIOR APPLICATION NUMBER: 09/016,434
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 1490
; SOFTWARE: PERL Program
SEQ ID NO 417
; LENGTH: 169
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20040010136A1 2137838
US-10-305-720-417

Alignment Scores:
Pred. No.: 3.75      Length: 169
```

Score: 79.50 Matches: 19
Percent Similarity: 53.57% Conservative: 11
Best Local Similarity: 33.93% Mismatches: 25
Query Match: 3.88% Indels: 1
DB: 17 Gaps: 1

US-10-077-111-13 (1-384) x US-10-305-720-417 (1-169)

QY 69 Leu1aSerCySerThrValLeuTrpAsnThrGluAsnGlyGluMet 88
DB 3 CTCATCACAGCTTCTTCAGATGTAACCTTAAAGATCTTGAAGAAGGCTC 62
QY 89 Leu1aValMetGluGlnProSerGlySerProValArgValCysGlnPheSerProAsp 108
DB 63 ATCATATACACTTCAAGACATACCGGA---CCTGCTTACTAGTTTCAAAAAGT 119
QY 109 SerThrCySerLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrp 124
DB 120 GGAGAGCTATTGTGATCAGAGGTGCAGACACACAGGCTTATATATG 167

RESULT 8

US-10-029-386-16111/C
Sequence 16111, Application US/10029386
Publication No. US20030194704A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Hanzel, David K.
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
FILE REFERENCE: AROMICA-X-2
CURRENT APPLICATION NUMBER: US/10/029,386
CURRENT FILING DATE: 2001-12-20
NUMBER OF SEQ ID NOS: 34288
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 16111
LENGTH: 155
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO CHR14.1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.2
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.2
OTHER INFORMATION: SWISSPROT HIT: P38262, EVALUATE 5.90e-02
OTHER INFORMATION: NT HIT: g113649630, EVALUATE 8.00e-74
OTHER INFORMATION: EST_HUMAN HIT: BE739840.1, EVALUATE 1.00e-73
US-10-029-386-16111

Alignment Scores:

Pred. No.: 4.3 Length: 155
Score: 78.50 Matches: 19
Percent Similarity: 40.58% Conservative: 9
Best Local Similarity: 27.54% Mismatches: 18
Query Match: 3.83% Indels: 24
DB: 16 Gaps: 1

US-10-077-111-13 (1-384) x US-10-029-386-16111 (1-155)

QY 107 ProAspSerThrCySerLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrpAsnAla 126
DB 144 CCAGATGACACATATATACCAAGAGTTCTGCACCCGTCGCTGCGCTTG----- 91
QY 127 GlnSerThrValLeuThrArgCysGlySerValValAspGlySerLeuAlaAlaCysAla 146
DB 91 ----- 91
QY 147 PheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAspLeuThrValTrp 166

DB 90 ---GCACCAGATGTTCCATGACAGTATCTGGAATCAAGCTGGGCAACTATCTTGTGG 34
QY 167 AspAspSerMetArgCysLeuHisSer 175
DB 33 CAGGAAGCTAAGGC-TGTGGCCACAGC 8

RESULT 9

US-09-998-598-205/C
Sequence 205, Application US/09998598
Patent No. US20020150922A1
GENERAL INFORMATION:
APPLICANT: Stolk, John A.
APPLICANT: Xu, Jiangchun
APPLICANT: Chenault, Ruth A.
APPLICANT: Meagher, Madelein Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.561
CURRENT APPLICATION NUMBER: US/09/998,598
CURRENT FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 2606
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 205
LENGTH: 197
TYPE: DNA
ORGANISM: Homo sapiens
US-09-998-598-205

Alignment Scores:

Pred. No.: 32.6 Length: 197
Score: 72.50 Matches: 21
Percent Similarity: 44.62% Conservative: 8
Best Local Similarity: 32.31% Mismatches: 17
Query Match: 3.54% Indels: 19
DB: 9 Gaps: 3

US-10-077-111-13 (1-384) x US-09-998-598-205 (1-197)

QY 31 AspLeuThrIleArgLeu-----TyrSerLeuArgAspPhe 42
DB 194 GATTAACCATCAACAGTATGGAATACCTGGGTGTGCAAAATGACACTGTCAGAGAT--- 138
QY 43 ThrGluLeuProHisSerProLeuLeuPheHisThrValAlaValHisCysCysPhe 62
DB 137 -----GAGAGCCACTCAGAGTGGGTGCTTGTGTCGCTTC 102
QY 63 SerPro-----SerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
DB 101 TCGCCCAACAGCAGCAACCTATATCATGCTCTCTGCTGGCGGACAGAGCTGCAAGGTA 42
QY 81 TrpAsnThrGluAsn 85
DB 41 TGGAACTGGCTAAC 27

RESULT 10

US-09-998-598-2327
Sequence 2327, Application US/09998598
Patent No. US20020150922A1
GENERAL INFORMATION:
APPLICANT: Stolk, John A.
APPLICANT: Xu, Jiangchun
APPLICANT: Chenault, Ruth A.
APPLICANT: Meagher, Madelein Joy
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.561
CURRENT APPLICATION NUMBER: US/09/998,598
CURRENT FILING DATE: 2001-11-16
NUMBER OF SEQ ID NOS: 2606
SOFTWARE: Corixa Invention Disclosure Database
SEQ ID NO 2327
LENGTH: 197

```
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-598-598-2327

Alignment Scores:
Pred. No.: 32.6      Length: 197
Score: 72.50      Matches: 21
Percent Similarity: 44.62%      Conservative: 8
Best Local Similarity: 32.31%      Mismatches: 17
Query Match: 3.54%      Indels: 19
DB: 9      Gaps: 3

US-10-077-111-13 (1-384) x US-09-598-598-2327 (1-197)

QY 31 AppLySerThrIleArgLeu-----TyrSerLeuArgAspPhe 42
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4 GATTAACCATCAAGCATGATGATACCTGGGTGTGCAAAATCACTGCGAAGT--- 60

QY 43 ThrGluLeuProHisSerProLeuLysPheHisThrTyraIaValHisCysCysPhe 62
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 61 -----GAGAGCCACTCAGAGTGGGTGTCTGTGTCGCCCTTC 96

QY 63 SerPro-----SerGlyHisIleLeuAlaSerCysSerThrIleArgIleThrValLeu 80
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 TCGCCCAACAGCAGCAACCTATCATCTGCTCTGCTGGGACAGCAAGTGTCAAGTA 156

QY 81 TrpAsnThrGluAsn 85
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 TGGAACTGTGGCTAAC 171

RESULT 11
US-10-425-115-24663
/ Sequence 24663, Application US/10425115
/ Publication No. US20040214272A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ NUMBER OF SEQ ID NOS: 2003-04-28
/ SEQ ID NO 24663
/ LENGTH: 187
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_1224C.1
US-10-425-115-24663

Alignment Scores:
Pred. No.: 34.5      Length: 187
Score: 72.00      Matches: 14
Percent Similarity: 52.94%      Conservative: 4
Best Local Similarity: 41.18%      Mismatches: 16
Query Match: 3.52%      Indels: 0
DB: 18      Gaps: 0

US-10-077-111-13 (1-384) x US-10-425-115-24663 (1-187)

QY 345 TrpIleSerIysIysArgThrSerProMetThrAsnLeuValLeuProSerAlaVal 364
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2 TGGTTCCGCCCGCCGCAAGTGTCTGCCCAAGACCGGACGAGTCTGAATCTGGAG 61

QY 365 LeuThrProAsnArgThrLeuLysMetAlaIleAsnArgTrp 378
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 62 CTGTGTCCCAACAGGCGCTCAGAACTGATCTCCGGCTGG 103

RESULT 12
US-10-424-599-124934
```

```
/ Sequence 124934, Application US/10424599
/ Publication No. US20040031072A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J
/ APPLICANT: Kovacic, David K
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
/ FILE REFERENCE: 38-21(53223)B
/ CURRENT APPLICATION NUMBER: US/10/424,599
/ NUMBER OF SEQ ID NOS: 285684
/ SEQ ID NO 124934
/ LENGTH: 169
/ TYPE: DNA
/ ORGANISM: Glycine max
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT3847_83824C.1
US-10-424-599-124934

Alignment Scores:
Pred. No.: 33.7      Length: 169
Score: 71.50      Matches: 15
Percent Similarity: 50.00%      Conservative: 9
Best Local Similarity: 31.25%      Mismatches: 17
Query Match: 3.49%      Indels: 7
DB: 17      Gaps: 1

US-10-077-111-13 (1-384) x US-10-424-599-124934 (1-169)

QY 125 AsnAlaGlnSerTyrlsLeuTyraArgCysGlySerValIysAspGlySerLeuAlaIa 144
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 23 AATGCAATGCCAACAATCTTATCGAGTATCCCAACACCTCTGATGCAATATCTAGT 82

QY 145 CysAlaPheSerProAsnGlySerPheIleValThrGlySerCysGlyAspLeuThr 164
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 83 ATTAGCTCAGTCCCAAGCAATGCTTCTGCACACT-----ACT 121

QY 165 ValTrpAspAspIysMetArgCys 172
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 122 TCATGGGACACCAAGTTCGGTGT 145

RESULT 13
US-10-029-386-25501/c
/ Sequence 25501, Application US/10029386
/ Publication No. US20030194704A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharron G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: AEOMICA-X-2
/ CURRENT APPLICATION NUMBER: US/10/029,386
/ NUMBER OF SEQ ID NOS: 2001-12-20
/ SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 25501
/ LENGTH: 131
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO CHR19.3
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.8
/ OTHER INFORMATION: SWISSPROT HIT: P90587, EVALUAE 4.00e-04
/ OTHER INFORMATION: EST HUMAN HIT: B1560259.1, EVALUAE 5.00e-23
/ OTHER INFORMATION: NT HIT: AF072879.1, EVALUAE 4.40e-01
US-10-029-386-25501
```

Alignment Scores:

| Pred. No.: | 67.4 | Length: | 131 |
|------------------------|--------|---------------|-----|
| Score: | 67.50 | Matches: | 13 |
| Percent Similarity: | 50.00% | Conservative: | 7 |
| Best Local Similarity: | 32.50% | Mismatches: | 19 |
| Query Match: | 3.30% | Indels: | 1 |
| DB: | 16 | Gaps: | 1 |

US-10-077-111-13 (1-384) x US-10-029-386-25501 (1-131)

QY 53 HistHyrAlaValAlaHisCysGlyCysPheSerProSerGlyHisLeuAlaSerCys 72
| | | | | : : : : : | | | | | : : : : :
DB 125 CACACAGAGGTCATGCATGCTGCTTGGACCCGACAGCCAGAGGCTGCTTGTCT 66
| | | | | : : : : : | | | | | : : : : :
QY 73 SerThrAspGlyThrThrValLeuTrpAsn---ThrGluAsnGlyGlnMetLeuAlaVal 91
| | | | | : : : : : | | | | | : : : : :
DB 65 TCATTGGACAGGTCATCAAGATCTGGAGTGTATACATCCAGCCAGCCAGCTGCTCACCATC 6

RESULT 14

US-10-060-036-1810
; Sequence 1810, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060.036
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1810
; LENGTH: 167
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 111, 145
; OTHER INFORMATION: n = A,T,C or G
US-10-060-036-1810

Alignment Scores:

| Pred. No.: | 171 | Length: | 167 |
|------------------------|--------|---------------|-----|
| Score: | 65.50 | Matches: | 17 |
| Percent Similarity: | 42.00% | Conservative: | 4 |
| Best Local Similarity: | 34.00% | Mismatches: | 28 |
| Query Match: | 3.20% | Indels: | 1 |
| DB: | 14 | Gaps: | 1 |

US-10-077-111-13 (1-384) x US-10-060-036-1810 (1-167)

QY 69 LeuAlaSerCysSerThrAspGlyThrThrValLeuTrpAsnThrGluAsnGlyGlnMet 88
| | | | | : : : : : | | | | | : : : : :
DB 14 ATATCATCCACGCTCTGGGATACCACTGCTGCGGACATTTGAGACAGGCCACAG 73
| | | | | : : : : : | | | | | : : : : :
QY 89 LeuAlaValMetGluGlnProSerGlySerProValArgValCysGlnPheSerProAsp 108
| | | | | : : : : : | | | | | : : : : :
DB 74 ACAAGTGGGTTTGTCTGACACAGTGGGAT---GTGATGTCCTGCTGCGCCCGAT 130
| | | | | : : : : : | | | | | : : : : :
QY 109 SerThrCysLeuAlaSerGlyAlaAlaAsp 118
| | | | | : : : : : | | | | | : : : : :
DB 131 GCGCGACAGCTTTGTTGTCAGGCGCCTGTGAT 160
| | | | | : : : : : | | | | | : : : : :
RESULT 15

US-09-864-761-23956/c
; Sequence 23956, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Neomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864.761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 23956
; LENGTH: 131
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC019221.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 5
; OTHER INFORMATION: SWISSPROT HIT: Q16760, EVALU6 7.00e-21
; OTHER INFORMATION: NT HIT: D63479.2, EVALU6 4.00e-68
; OTHER INFORMATION: EST_HUMAN HIT: BB88147.1, EVALU6 6.00e-68
US-09-864-761-23956

Alignment Scores:

| Pred. No.: | 134 | Length: | 131 |
|------------------------|--------|---------------|-----|
| Score: | 65.00 | Matches: | 15 |
| Percent Similarity: | 52.50% | Conservative: | 6 |
| Best Local Similarity: | 37.50% | Mismatches: | 19 |
| Query Match: | 3.18% | Indels: | 0 |
| DB: | 9 | Gaps: | 0 |

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 5, 2005, 11:55:52 ; Search time 694 Seconds

(without alignments)
3184.733 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047

Sequence: 1 MYXLIHTLADHGDVNCACF.....LTPNRLKMAINRWLETHQK 384

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapext 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 4313806 segs, 2877871033 residues

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n.model -DEV=x1h

-O=/cg2_1/USPTO.spool/US10077111/runat_04022005_121054_6859/app.query.fasta.1.583

-DB=Published.Applications.NA -QFMT=fastap -SUFFIX=g2n.rmpb -MINMATCH=0.1

-LOOPCL=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=blomsum62

-TRANS=human40.cdt -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100

-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0

-MALEN=2000000000 -USER=US10077111 @CGN_1.1.628 @runat_04022005_121054_6859

-NCPU=6 -ICPU=3 -NO MMAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100

-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5

-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:*

1: /cg2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*

2: /cg2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*

3: /cg2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*

4: /cg2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*

5: /cg2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*

6: /cg2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*

7: /cg2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*

8: /cg2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*

9: /cg2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*

10: /cg2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*

11: /cg2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*

12: /cg2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*

13: /cg2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*

14: /cg2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*

15: /cg2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*

16: /cg2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*

17: /cg2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*

18: /cg2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*

19: /cg2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*

20: /cg2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*

21: /cg2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*

22: /cg2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|----------------------|
| 1 | 2047 | 100.0 | 1272 | 13 | US-10-077-111-12 |
| 2 | 1984 | 96.9 | 1553 | 13 | US-10-077-111-3 |
| 3 | 1984 | 96.9 | 1818 | 13 | US-10-077-111-1 |
| 4 | 1984 | 96.9 | 1996 | 18 | US-10-357-930-23179 |
| 5 | 1984 | 96.9 | 1996 | 18 | US-10-357-930-29046 |
| 6 | 1978 | 96.6 | 1773 | 16 | US-10-287-218-40 |
| 7 | 1978 | 96.6 | 1773 | 18 | US-10-474-291-40 |
| 8 | 1978 | 96.6 | 1811 | 17 | US-10-104-047-1689 |
| 9 | 1978 | 96.6 | 1817 | 10 | US-09-971-392-207 |
| 10 | 1978 | 96.6 | 1844 | 15 | US-10-037-270-768 |
| 11 | 1978 | 96.6 | 1844 | 17 | US-10-117-722-768 |
| 12 | 1958 | 95.7 | 1908 | 13 | US-10-077-111-14 |
| 13 | 1616.5 | 79.0 | 1901 | 13 | US-10-077-111-6 |
| 14 | 709 | 34.6 | 630 | 13 | US-10-077-111-5 |
| 15 | 627.5 | 30.7 | 446 | 18 | US-10-357-930-5029 |
| 16 | 626.5 | 30.6 | 409 | 18 | US-10-357-930-14198 |
| 17 | 621.5 | 30.4 | 441 | 18 | US-10-357-930-35319 |
| 18 | 621.5 | 30.4 | 441 | 18 | US-10-357-930-44154 |
| 19 | 597 | 29.2 | 366 | 9 | US-09-864-761-10657 |
| 20 | 532 | 26.0 | 297 | 9 | US-09-864-761-27306 |
| 21 | 352 | 17.2 | 520 | 13 | US-10-077-111-8 |
| 22 | 308.5 | 15.1 | 466 | 9 | US-09-864-761-11052 |
| 23 | 300 | 14.7 | 180 | 9 | US-09-864-761-27687 |
| 24 | 264.5 | 12.9 | 837 | 17 | US-10-369-493-43585 |
| 25 | 264 | 12.9 | 3015 | 17 | US-10-369-493-43529 |
| 26 | 263.5 | 12.9 | 1593 | 17 | US-10-369-493-36600 |
| 27 | 259.5 | 12.7 | 3025 | 9 | US-09-764-853-29 |
| 28 | 259 | 12.7 | 3082 | 17 | US-10-451-207-18 |
| 29 | 256.5 | 12.5 | 3166 | 17 | US-10-451-207-13 |
| 30 | 251 | 12.3 | 1137 | 17 | US-10-369-493-31840 |
| 31 | 250 | 12.2 | 1741 | 17 | US-10-369-493-42580 |
| 32 | 246.5 | 12.0 | 1830 | 17 | US-10-369-493-43872 |
| 33 | 245 | 12.0 | 840 | 17 | US-10-369-493-43586 |
| 34 | 244.5 | 11.9 | 3420 | 17 | US-10-369-493-42740 |
| 35 | 242.5 | 11.8 | 1965 | 17 | US-10-369-493-42753 |
| 36 | 242.5 | 11.8 | 3408 | 17 | US-10-369-493-42733 |
| 37 | 236.5 | 11.6 | 3358 | 17 | US-10-369-493-42532 |
| 38 | 232.5 | 11.4 | 3465 | 17 | US-10-369-493-43556 |
| 39 | 231.5 | 11.3 | 1818 | 17 | US-10-369-493-42594 |
| 40 | 230 | 11.2 | 852 | 17 | US-10-369-493-42751 |
| 41 | 230 | 11.2 | 2768 | 17 | US-10-424-599-115833 |
| 42 | 229.5 | 11.2 | 852 | 17 | US-10-369-493-42703 |
| 43 | 229 | 11.2 | 480 | 9 | US-09-864-761-13978 |
| 44 | 229 | 11.2 | 1434 | 17 | US-10-369-493-42592 |
| 45 | 227 | 11.1 | 7042 | 9 | US-09-876-667-1 |

ALIGNMENTS

RESULT 1
US-10-077-111-12
Sequence 12, Application US/10077111
Publication No. US20020187492A1
GENERAL INFORMATION:
APPLICANT: Toddlerud, C. Gordon
APPLICANT: Finger, Joshua N.
APPLICANT: Rillema, Jijia
TITLE OR INVENTION: TBA
FILE REFERENCE: 3053-411AUS2
CURRENT APPLICATION NUMBER: US/10/077, 111
CURRENT FILING DATE: 2002-02-15
PRIOR APPLICATION NUMBER: 60/294,181
PRIOR FILING DATE: 2001-05-29
PRIOR APPLICATION NUMBER: 60/269,366
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 12
LENGTH: 1272


```

; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: RET 16.2 splice variant
US-10-077-111-12

```

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|-----|
| Pred. No.: | 2,74e-233 | Length: | 127 |
| Score: | 100.00 | Matches: | 364 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 100.00% | Indels: | 0 |
| DB: | 13 | Gaps: | 0 |

US-10-077-111-13 (1-384) x US-10-077-111-12 (1-1272)

| | | | |
|----|-----|---|------|
| OY | 1 | MeGVaLVyLsLeu11GhiEThrLeuAlaAPh1sGLYAspAspValAsnCYsValaPhe | 20 |
| Db | 111 | ATGGGAAACTGCATTCCACCATTAAGCTGATCATTGAGACGATCTCACTGCTGTGCTTC | 170 |
| OY | 21 | SerPheSerLeuLeuAlaThrCYsSerLeuAspLYsThr1LeuLeuTYsSerLeuArg | 40 |
| Db | 171 | TCCCTTTCCCTCTTGCGTACTTGCTCTTCGGGCAAAACAACTTCCGCTGTAAGCT | 230 |
| OY | 41 | AspPheThrGluLeuProHisSerProLeuLYsPheHisEThrTYrValaHisCYsCYs | 60 |
| Db | 231 | GACTTACTGAACTGCGCACTTCTCCATTGAACTTTCATACCTAAGCTGTCCACTGCTGC | 290 |
| OY | 61 | CysPheSerProSerGlyHis1LeuAlaSerCYsSerThrAspGlyThrValLeu | 80 |
| Db | 291 | TGTTTCTCCCTTCAGACATATTTTGGCATCGTGTCAACAGATGTAACCATGTGCTCTTA | 350 |
| OY | 81 | ThrPantThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal | 100 |
| Db | 351 | TGGAATACTGAAATGGAACAGATGCTGGGACGATGGAACAGCTGATGGCACCTGTG | 410 |
| OY | 101 | ArgValCYsGlnPheSerProAspSerThrCYsLeuAlaSerGlyVala1aAspGlyThr | 120 |
| Db | 411 | AGGCTTGCCACTTTTCCCACTCCAGTGTGTTGGCATCAGGGGCACTGATGGAAGCT | 470 |
| OY | 121 | ValValLeuThrPAsnAlaGlnSerTYrLYsLeuTYrArgCYsGlySerValLYsAspGly | 140 |
| Db | 471 | GTCGCTTTTGTGGAAAGCAGCATCATCAATTAATATAGATGTGTGATGTTAAAGATGCG | 530 |
| OY | 141 | SerLeuAlaAlaCYsAlaPheSerProAsnGlySerPhePheAlaThrGlySerSerCYs | 160 |
| Db | 531 | TCCTTGCGGCGCATGTGCATTTTCTCCTAATGGAAGCTTCTTGTCACTGGCTCTCATGT | 590 |
| OY | 161 | GlyAspLeuThrVal1ThrAspAspLYsMetArgCYsLeuHisSerGlyLYsAlaHisAsp | 180 |
| Db | 591 | GGTGATTTTAACGTGTGGGATGATTAATAGAGGTGTCTGCATGTGTAAGCAACATGAT | 650 |
| OY | 181 | LeuGly11LeThrCYsCYsAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu | 200 |
| Db | 651 | CTTGGAATTAACCTGCTGCGCATTTTCTTCAACAGCCAGTTCTGATGGAGAACAAAGCTTT | 710 |
| OY | 201 | GlnPhePheArgLeuAlaSerCYsGlyGlnAspCYsGlnValLYs1LeThrP1LeValSer | 220 |
| Db | 711 | CAGTTTTTCAGCTGCGATCATGTGTGTCAGGATGTCCAAGTCCAAATTTTGGATTGTTCT | 770 |
| OY | 221 | PheThrHis1LeLeuAlaArgArgThrGluHisGlnLeuLYsGlnPheThrGluAspTrp | 240 |
| Db | 771 | TTTACCCCATATCTTTCAGCAAGCGCACAGAACATCACTGAAAGCAATTTACCGAAGATTGG | 830 |
| OY | 241 | SerGluGluValValSerThrTrpLeuCYsAlaGlnAspLeuLYsAspLeuValGly1Le | 260 |
| Db | 831 | TCAGAGGAGGTGCTTCACATGCGCTTGTGTGCACAAAGATTTAAAGATCTTGTGGATTT | 890 |
| OY | 261 | PheLYsMetAsnAsn1LeAspGlyLYsGluLeuLeuAsnLeuThrArgGlnSerLeuAla | 280 |
| Db | 891 | TTTCAAGATGAAATACATCTTATGGAAGAAAGACTGTTGATCTTTCAAAGAAAGACTGTGGCT | 950 |
| OY | 281 | AspAspLeuLYs1LeGlnSerLeuGlyLeuArgSerLYsValLeuArgLYs1LeGluGlu | 3000 |

| | | | |
|----|------|---|------|
| Db | 951 | GATATTTGAAAATTTGAAATCTCTAGAGCTCGTAGTAAGTGGTAGGAAAATTGAAGAG | 1010 |
| Qy | 301 | LeuArgThrIlyValIlySerLeuSerSerGlyIleProAspGluPheIleCysProIle | 320 |
| Db | 1011 | CTCAGAGACCAAGGTTAAATCCCTTCTTCAGGAATTCGTATGAATTTAATATGTCACATA | 1070 |
| Qy | 321 | ThrArgGluLeuMetIyAspProValIleIleSerAspGlyIyrSerTyrGluIyGlu | 340 |
| Db | 1071 | ACTAGAGAACTTAGAAAAGATCCCGGTATGCAATGCATGATGGCTATTCATATGAAAAGAA | 1130 |
| Qy | 341 | AlaMetGluAsnTrpIleSerIlyValIyArgThrSerProMetThrAsnLeuValIleu | 360 |
| Db | 1131 | GCATGTGAAAATTGATCAGCAAAAAGAAACGTACAAATGCCATGACAAATCTTGTTCTT | 1190 |
| Qy | 361 | ProSerAlaValLeuThrProAsnArgThrLeuIyMetAlaIleAsnArgTrpLeuGlu | 380 |
| Db | 1191 | CCTTCACCGGTACTTACACCAATATGAGACTCTGAAATGGCCATCAATATAGATGGCTGAG | 1250 |
| Qy | 381 | ThrIleGluIyS 384 | |
| Db | 1251 | ACACACCAAAAG 1262 | |

RESULT 2

```

US-10-077-111-3
: Sequence 3, Application US/10077111
: Publication No. US2002018792A1
: GENERAL INFORMATION:
: APPLICANT: Toddertud, C. Gordon
: APPLICANT: Finger, Joshua N.
: APPLICANT: Rillema, Jill
: TITLE OF INVENTION: TBA
: FILE REFERENCE: 3053-41AUS2
: CURRENT APPLICATION NUMBER: US/10/077,111
: CURRENT FILING DATE: 2002-02-15
: PRIOR APPLICATION NUMBER: 60/294,181
: PRIOR FILING DATE: 2001-05-29
: PRIOR APPLICATION NUMBER: 60/269,366
: PRIOR FILING DATE: 2001-02-16
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn Ver. 2.1

```

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|---------|
| Pred. No.: | 1,22e-225 | Length: | 155,327 |
| Score: | 1984.40 | Matches: | 383 |
| Percent Similarity: | 80.46 | Conservative: | 0 |
| Best Local Similarity: | 80.46 | Mismatches: | 1 |
| Query Match: | 96.92 | Indels: | 92 |
| DB: | 13 | Gaps: | 1 |

US-10-077-111-13 (1-384) X US-10-077-111-3 (1-1553)

| | | | |
|----|-----|---|-----|
| OY | 1 | MetValValLeuIleIleIleThrlleuAlaAspHisgLyAspAspValAsnGlyCysAlaIlePhe | 20 |
| Db | 111 | ATGGGAAACTGATTCACACATTACCTGATCATCGTGAACGATGTCAACCTGCTGCTTC | 170 |
| OY | 21 | SerPheSerLeuLeuAlaThrCysSerLeuAspLysIleIleArgLeuTyrSerLeuArg | 40 |
| Db | 171 | TCCCTTTTCCCTTGGCTACCTTGGCTCCTTGGACAAACAATTCGGCTGTAACCTGGTAAGT | 230 |
| OY | 41 | AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys | 60 |
| Db | 231 | GACTTAACTGAACTCCACATTCCTCATTTCAATGAAATTCACTACCTAGCTGCCACTGCTGC | 290 |
| OY | 61 | CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu | 80 |
| Db | 291 | TGTTTCTCCCTTACGAGCATATTTTGGCATCTGTTCACAGAGATGGTCACTGTGCTCCA | 350 |

QY 81 TRPAsnThrGluuansnglyglMetLeuAlaValMetGluInProSerGlySerProVal 100
DB 351 TGGAAATACGAAAAATGACAGATGCTGGCAGTGAACAGCCTAGTGGCAGCCCTGTG 410
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
DB 411 AGGGATTGCGAGTTTCCCGAGACTCCAGCTGTTGGCATCAGGGGAGCGTGAAGACT 470
QY 121 ValValLeuThrPasnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly 140
DB 471 GTGGTTTGTGGAATGACAGATCATACAAATTATATAGATGTGATGTTAAAGATGC 530
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
DB 531 TCCTGGCGGCGCATGTGATTTCTCCAAATGGAAGCTTCTTGTCACCTGCTCCTCATGT 590
QY 161 GlyAspLeuThrValTyrPasnAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
DB 591 GGTGATTTAACAGTGTGGATGATMAAATAGGTGTGATGATGAAAAAGACATGAT 650
QY 181 LeuGlyLysThrCysAspPheSerSerSerGlnProValSerAspGlyGlnGlyLeu 200
DB 651 CTGGAAATTAACGCTGCGATTTTCTTTCACAGCAGTTTCTGATGAGAACAGATCTT 710
QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysLysLysLysLysLys 220
DB 711 CAGTTTCTTTCGACCTGGCATGATGATGATGATGATGATGATGATGATGATGATGAT 770
QY 221 PheThrHisGlyLeu----- 225
DB 771 TTTACCATATCTTAGGTTTGAATTAAATATATAAGTACACTGAGTGGCAGCTGTCT 830
QY 225 ----- 225
DB 831 CCTGTTCTGCTGCTGCTGCTTTCCTCCGATGAGGACAGATGCTAGTCTCAGGGTCAATGAT 890
QY 225 ----- 225
DB 891 AAGTGTGTCAATAGATATGATACTAATATAGAGATATAGTCAACATGATGATGATGAT 950
QY 225 ----- 225
DB 951 ACCAGTATGTCACAACTTGCTGCTTTCGACCTAATACCTTTTACTGCTAGTGTTC 1010
QY 226 -----AlaArgArg 228
DB 1011 ATGGAACAAAACAGTAAACATCTGGCAATTTGACCTGGAAACACTTTGCCAAGAGGCGC 1070
QY 229 ThrGlnHisGlnLeuLysGlnPheThrGlnAspTyrSerGlyGlnValLysSerThrTyr 248
DB 1071 ACAGAACATCAGCTGAAGCAATTTACCGAAGATTTGGTCAGAGAGATGTCTCAACATGG 1130
QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyLysPheLysMetAsnAsnLysAspGly 268
DB 1131 CTGTGTCACAAAGATTTAAAGATCTGTGTAATTTTCAAGATGAATTAACATGATGATGA 1190
QY 269 LysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysLysGlnSerLeu 288
DB 1191 AAGGAACATGTTGAATCTTCAAAAAGAAAGTCTGCGATGATTTGAAAATTTGAATCTCTA 1250
QY 289 GlyLeuArgSerLysValLeuArgLysLysGlnGlnLeuArgThrLysValLysSerLeu 308
DB 1251 GGAAGTGTGTAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1310
QY 309 SerSerGlyLysProAspGlnPheLysCysProLysThrArgGlnLeuMetLysAspPro 328
DB 1311 TCTTTCAGAAATCTCTGATGATTTATATGTCATTAATCTAGAGAACTTAAGAAAGATCCG 1370
QY 329 ValLysLeuSerAspGlyTyrSerTyrGlnLysGlnAlaMetGlnAsnThrLysSerLys 348
DB 1371 GTATTCGCATCAAGTGGCTATTCATATGAAGAAAGAACCAATGGAATTTGATCAGGAAA 1430
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368

DB 1431 AAGAAACGTACAGATCCCATGACAAATCTTGTTCTTTCACCGGTACTTACACCAAAAT 1490
QY 369 ArgThrLeuLysMetAlaLysAsnArgTyrPheGluThrHisGlnLys 384
DB 1491 AGGACTGTGAAAATGGCCATCATATAGATGGCTGGAGACACACCAAAAG 1538

RESULT 3
US-10-077-111-1
; Sequence 1, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Toddertud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jill
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; SOFTWARE: PatentIn Ver. 2.1
; NUMBER OF SEQ ID NOS: 25
; SEQ ID NO 1
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: HUMAN
US-10-077-111-1

Alignment Scores:
Pred. No.: 1 566-225 Length: 1818
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 80.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: 13 Gaps: 1

US-10-077-111-13 (1-384) x US-10-077-111-1 (1-1818)

QY 1 MetValLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeuLysLeu 20
DB 148 ATGTGAAACTGATTCACACATTAAGTGAATGATGATGATGATGATGATGATGATGAT 207
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrLysArgLeuTyrSerLeuArg 40
DB 208 TCCTTTTCCCTCTTGCTGCTAATCTGCTCTTGGACAAAACAAATTCGCTGATCGTTACGT 267
QY 41 AspPheThrGlnLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys 60
DB 268 GACTTACTGAACTGACACATTCCTCATTTGAAGTTTCACTATACCTATCTGCTGCTGCTG 327
QY 61 CysPheSerProSerGlyHisLysLeuAlaSerCysSerThrAspGlyThrValLysLeu 80
DB 328 TGTTTCTCCCTTCAGACATATTTTGGCATCGTTTCAACAGATGATGATGATGATGATGAT 387
QY 81 TrpAsnThrGluuansnglyglMetLeuAlaValMetGluInProSerGlySerProVal 100
DB 388 TGGAAATACGAAAAATGACAGATGCTGGCAGTATGAAACAGCTGATGCGACCCCTGTG 447
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
DB 448 AGGGATTGCGAGTTTCCCGAGCTCCAGCTGATGTTGGCATCAGGGGAGCGTGAAGACT 507
QY 121 ValValLeuThrPasnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly 140
DB 508 GTGGTTTGTGGAATGACAGATCATACAAATTATATAGATGTGATGTTAAAGATGCG 567
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
DB 568 TCCTTGGCGGCGCATGTGATTTCTCCTAAATGGAAGCTTCTTGTCACCTGCTCCTCATGT 627

QY 161 G1yAspLeuThrVal1TrpAspAspLysMetArgCysLeuHisSerGluLysAlaHisAsp 180
496 GGTATTTTAAACAGTGGGATGATTAATATGAGTGTCTGCATATGAAAAAGCAATGAT 555
QY 181 LeuG1y1LeThrCysCysAspPheSerSerGlnProValSerAspGlyValGluGlyLeu 200
556 CTTGGAAATTACCTGCTCGATTTTCTTCAACAGCCAGTTTCTGATGAGAAACAAGTCTT 615
QY 201 G1hPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLys1LeThrPheValSer 220
616 CAGTTTTCGACTGGCAATCATGTGTCAGGATGGCCAAAGTCAAAATTTGATTTGTTCT 675
QY 221 PheThrHis1LeuLeu----- 225
676 TTTTACCATATCTTAGGTTTGAATTAATAATATAAGTACACTGAGTGGCACTGTGCT 735
QY 225 ----- 225
736 CTTGTCTGGCTGTGCTTTTCCATGATGGCAGATGCTAGTCAAGGTGATGAT 795
QY 225 ----- 225
796 AAGTCTGTCAATGATATGATATGATACTAATACTGAGAAATATCTTCAACATGACTCAGCAC 855
QY 225 ----- 225
856 ACCAGATATGTCACAACTTGCTTTTGCACTTAATACCTTTTACTTGTCTACTGTCTTCA 915
QY 226 -----AlaArgArg 228
916 ATGGACAATAACAGTGAACATCTGGCAATTGACCTGGAAACACTTGGCAAGAGGCGC 975
QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValValSerThrTrp 248
976 ACAGAAATCATGAGTGAAGCAATTTTACGAAATTTGTCAGAGGAGATGCTCTCAACATGG 1035
QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGly1LePheLysMetLysAsn1LeAspGly 268
1036 CTTTGTGCACAAAGTTTAAAGATCTTGTGTATTTTTCAGATGAAATTAACATTTGATGGA 1095
QY 269 LysGluLeuLeuLeuLeuThrLysGluSerLeuAlaAspPheLys1LeGluSerLeu 288
1096 AAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGTTGAAATTTGAATCTCTTA 1155
QY 289 GlyLeuArgSerLysValLeuArgLys1LeGluLeuLeuArgThrLysValLysSerLeu 308
1156 GGACTGGTGTAGTAAAGTCTGAGGAAATGCAAGACTCAGAGACCAAGGTTAAATCCCTT 1215
QY 309 SerSerGly1LeProAspGluPhe1LeCysPro1LeThrArgGluLeuMetLysAspPro 328
1216 TCTTCAGGAATCTCTGATGAAATTTATATGTCATTAATCTAGCAAACTTATGAAAGATCCG 1275
QY 329 Val1LeAlaSerAspGly1TrpSerLysGluLysGluAlaMetGluAsnTrp1LeSerLys 348
1276 GTATGCCATCAGATGGCTATTTCAATGAAAGAAAGAAATGGAATTCAGCAAA 1335
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
1336 AAGAAACGTACAAATCCCATGACAAATCTTGTCTTCCATCAGCGGTATCTTACCACAAT 1395
QY 369 ArgThrLeuLysMetAla1LeAsnArgTrpLeuGluThrHisGlnLys 384
1396 AGGACTCTGAATAATGCGCATATGATGCTGAGACACACCAAAAG 1443
Db

RESULT 5
US-10-357-930-29046
; Sequence 29046, Application US/10357930
; Publication No. US20040255086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John

; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ. ID NOS: 62232
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 29046
; LENGTH: 1996
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994,
; LOCATION: 1995, 1996
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29046

Alignment Scores:

Pred. No.: 1,81e-225 Length: 1996
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 96.92% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: 18 Gaps: 1

US-10-077-111-13 (1-384) x US-10-357-930-29046 (1-1996)

QY 1 MetValLysLeuLeuHisThrLeuAlaAspHisGlyAspAspValaAspCysValaPhe 20
16 ATGTTGAATCTGATTTCAACATTTAGCTGATCATGATGATGATGATGATGATGATGATGAT 75
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThr1LeArgLeuTrpSerLeuArg 40
76 TCCCTTTCCCTCTTGGCTATTTGCTGCTTGGACAAAACAATTCGCTGATGATGATGATGAT 135
QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTrpAlaValHisCysCys 60
136 GACTTTACTTAACCTGACCACTTCTCCATTAAGTTTCATACATGCTGTCACCTGCTGC 195
QY 61 CysPheSerProSerGlyHis1LeLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
196 TGTTCCTCCCTTCAGAGCAATTTTGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 255
QY 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
256 TGAATATCTGAATAATGACAGATGCTGCACTGATGATGATGATGATGATGATGATGATGAT 315
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
316 AGGTTTGCAGATTTTCCAGACTCCAAGTGTGGCATCAGGGGCAAGCTGATGAACT 375
QY 121 ValValLeuThrPheAlaGlnSerTrpLysLeuTrpArgCysGlySerValLysAspGly 140
376 GTGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 435
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160

```
Db 436 TCCTGGCGGCGATGCGATTTTCTCCTAATGAGACTTCTTTGTCACTGGCTCCATGCT 495
Qy 161 G1YAspLeuThrValTTPAspAplysMeLarGySeLueH1sSerG1uLyLaH1sAsp 180
Db 496 GGGGATTTAACAGTGTGGGATGATTAATAAGGTGTCTGATATGTAATAAAGACATGAT 555
Qy 181 LeuG1Y1LeThrCySeCyAspPheSerSerG1nProValSerApg1yG1uG1nG1yLeu 200
Db 556 CTTGGAAATTACCGCTGCGGATTTTCTTTCACAGCCAGCTTCTGATGAGAACAAAGGCTT 615
Qy 201 G1nPhPheArgLeuA1aSerCySeG1yG1nApgCyG1nValy1eTTP1LeV1aSer 220
Db 616 CAGTTTTTTCAGATGCGCATCATGTGG1CAGAGATTGCGCAAGTCAAAATTGGATTGTTCT 675
Qy 221 PheThrH1e1Leu----- 225
Db 676 TTTTACCATATCTTAGGTTTGAATTAATAATAAGTACACTGAGTGGGCACTGTGCT 735
Qy 225 ----- 225
Db 736 CCTGTTCTGGCTTGTGCTTTTCCATGATGCGCAGATGCTAGTCTCAGGGTCAGTGAAT 795
Qy 225 ----- 225
Db 796 AAGTCTGTCAATGATATGATTAATTAATCTGAGAAATTAATCTTCAACATTTGACTCAGAC 855
Qy 225 ----- 225
Db 856 ACCAGTATGTCACACTGTGTGCTTTTGCACTAATACCTTTTACTTGCTACTGGTTCA 915
Qy 226 -----A1aATgATg 228
Db 916 ATGACAAAACAGTGAACATCTGCAATTGACCTGGAAAACCTTTGCCAAGCAAGCGCG 975
Qy 229 ThG1uH1eG1nLeuLySeG1nPhThrG1uApgTTPSerG1uG1uVal1aSerTTP 248
Db 976 ACAGAACTACAGTGAAGCAATTTCAGGAAGATTGG1CAGAGAGAGATGTCTCAACATGG 1035
Qy 249 LeuCyA1aG1nAAspLeuLyAspLeuVal1aG1Y1ePheLySeMeLarAen1aApg1y 268
Db 1036 CTTTGTGCAACAGATTTAAAGATCTTGTGGTATTTTTCAGATGAATTAACATTTGATGCA 1095
Qy 269 LySeG1uLeuAenLeuThrLySeG1uSerLeuA1aAspAAspLeuLy1eG1uSerLeu 288
Db 1096 AAAAGAACTGTTGAATCTTACAAAAGAAAGTGGCTGATGATTTGAAAATTTGAATCTCTA 1155
Qy 289 G1YLeuAArgSerLySValLeuAArgLyS1eG1uG1nLeuAArgThrLySVal1ySerLeu 308
Db 1156 GGACTGCGTAGTAAAGTGTGAGGAAAATTGAAGAGCTCAGACCAAGGTTAAATCCCTT 1215
Qy 309 SerSerG1Y1eProAAspG1uPhe1LeCyAPro1eThraG1uLeuMetLyAspPro 328
Db 1216 TCTTCAGAAATCTCTGATGAATTAATATGTCCAAATACTAGAACTTAAGAAAGATCCG 1275
Qy 329 Val1eA1aSerAenApg1yTTPSerLyG1uLySeG1uA1aMetG1uAenTTP1eSerLyS 348
Db 1276 GTCATGCGATCAGATGCTATTCATATGAAGAAAGCAATGAAATTTGATCAGAGAA 1335
Qy 349 LySeLyAArgThrSerProMetThraAenLeuValLeuProSerA1aVal1eThrProAen 368
Db 1336 AAGAAACGTAACAAGTCCCATGACAAATCTTGTCTTCCACGGTAATCTTACACCAAAAT 1395
Qy 369 ArgThrLeuLySeMeLarA1aAenAArgTTPLeuG1uThrH1eG1nLyS 384
Db 1396 AGGACTCTGAATAATGGCCATCAATAGTGGCTGAGAGACACCAAAAG 1443
```

```
RESULT 6
US-10-287-218-40
; Sequence 40, Application US//10287218
; Publication No. US20030198975A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.
; APPLICANT: AZIMZAI, Yalda; AU-YOUNG, Janice K.
```

```
; APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
; APPLICANT: BECHA, Shanya D.; BOROWSKY, Mark L.
; APPLICANT: BUFORD, Neill; DING, Li
; APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
; APPLICANT: GANDHI, Ameena R.; GIETZEN, Kimberly J.
; APPLICANT: GRIFPIN, Jennifer A.; HAFALIA, April J.A.
; APPLICANT: HONCHELU, Cynthia D.; LAL, Preeti G.
; APPLICANT: LEE, Soo Yeun; LU, Dyung Alina M.
; APPLICANT: ARVIZU, Chandre S.; RAMKUMAR, Jayalaxmi
; APPLICANT: REDDY, Roopa; SANJANWALA, Madhu, M.
; APPLICANT: TANG, Y. Tom; WALIA, Narinder K.
; APPLICANT: WANG, Yu-mei, E.; WARREN, Bridget A.
; APPLICANT: XU, Yuming; YANG, Junming
; APPLICANT: YAO, Monique G.; YUE, Henry
; APPLICANT: ZEBARJADIAN, Yeganeh
; TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
; FILE REFERENCE: PI-0417 USA
; CURRENT APPLICATION NUMBER: US/10/287,218
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: PCT/US02/11152
; PRIOR FILING DATE: 2002-04-05
; PRIOR APPLICATION NUMBER: US 60/349,705
; PRIOR FILING DATE: 2002-01-15
; PRIOR APPLICATION NUMBER: US 60/295,263
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/295,340
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/293,727
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: US 60/291,846
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/291,662
; PRIOR FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: US 60/287,228
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/286,820
; PRIOR FILING DATE: 2001-04-26
; PRIOR APPLICATION NUMBER: US 60/283,294
; PRIOR FILING DATE: 2001-04-11
; Remaining Prior Application data removed - See file Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: PERL Program
; SEQ ID NO 40
; LENGTH: 1773
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No. US20030198975A1 7483131CB1
US-10-287-218-40
Alignment Scores:
Pred. No.: 7,8e-225 Length: 1773
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 16 Gaps: 1
US-10-077-111-13 (1-384) x US-10-287-218-40 (1-1773)
Qy 1 MelValyLeuL1eH1eThrLeuA1aApgH1aG1YAspAAspValAAspCySeA1aPhe 20
Db 60 ATGGTAACCTGATTTACACATTAAGTATGATGATGATGATGATGATGATGATGATGATG 119
Qy 21 SerPheSerLeuLeuA1aThrCySeSerLeuAAspLySThr1LeAArgLeuTTPSerLeuAArg 40
Db 120 TCCTTTCCCTCTTGCTGCTACTTGTCTCTGAGCAAAACATTCGCTGTACTGTTACGT 179
Qy 41 AspPheThrG1uLeuProH1sSerProLeuLySePheH1eThrTTPA1aVal1aCySeCyS 60
Db 180 GACTTACTGAACTGCGCAATCTTCATTTGAAGTTTCAATCACTATGCTGCTGCTGCTGCTG 239
```

61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
240 TGTTCCTCCCTCCAGACATATTTGGCATCGTGTCAACAGATGGTACCTGCTCTA 299
81 ThrAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
300 TGGAAATACGAAATGACAGATGCTGGCAGTATGGAACAGCTTGAGCCCTGCTG 359
101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
360 AGGCTTGGCAGTTTCCCGACACTCCACTGTTGGCACAAGGCGAGCTGATGAACT 419
121 ValValLeuThrAsnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly 140
420 GTGCTTTTGGATGCACAGTCATACAAATTAATATGATGTGTGTTAAAGATGC 479
141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
480 TCTTGGCGGCAATGATGATTTCTCTTAATGAAGAGCTTCTTGTCACTGGCTCTCATGT 539
161 GlyAspLeuThrValTTPAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
540 GGTGATTTAACAGTGTGGATGATTAATAGGTGTCTGATGTGTAAGAAAGCAGCATGAT 599
181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlyLeu 200
600 CTGGATTAATCCTGCTGCGCATTTTCTTCAAGCAGATTTCTGATGGAAGAACAGGCTT 659
201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrrIleValSer 220
660 CAGTCTTTTGCAGCTGCATGATGTGTGTCAGATTTGCCAATGCAAAATTTGATTTCTT 719
221 PheThrHisIleLeu----- 225
720 TTTTACCATATCTTGAAGTTTGAATTAATAATTAAGTCACTGAGTGGCAGCTGTCT 779
225 ----- 225
780 CCGTTCTGGCTTGTGCTTTTCCCATGATGGGAGATGCTAGTCTGAGGGTCAATGAT 839
225 ----- 225
840 AAGTCTGTCAATGATATGATTAATTAATGAGATTAATTAATTAATTAATTAATTAAT 899
225 ----- 225
900 ACCAGTATGTCACAACTGTGCTTTTGCACCTAATACCTTTTACTGCTACTGTTCA 959
226 -----AlaArgArg 228
960 ATGACAAAACAGTGAACATCTGGCAATTTGACCTGGAACACTTTGCCAAGCAGAGAGC 1019
229 ThrGlnHisGlnLeuLysGlnPheThrGlnAspTyrSerGlnGlnValLysSerThrTyr 248
1020 ACAGAACATCAGCTGAAGCAATTTTACCGAAGATTTGTCAGAGAGAGATGCTCAACATG 1079
249 LeuCysAlaGlnPheLeuLysAspLeuValGlyIlePheLysMetLanMetIleAspGly 268
1080 CTTTGTGCAACAAATTTAAAGATCTTGTGTAATTTTCAAGATGAATTAACATTGATGGA 1139
269 LysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGlnSerLeu 288
1140 AAAGAACTGTGATCTTACAAAGAAAGCTGCTGATGATTTGAAATTTGATCTCTTA 1199
289 GlyLeuArgSerLysValLysArgLysIleGlnGlnLeuArgThrLysValLysSerLeu 308
1200 GGAAGTGGTGAAGAGTGAAGAAATGAAGAGCTCAGAGCAAGGTTAAATCCCTT 1259
309 SerSerGlyIleProAspGlnPheIleCysProIleThrArgGlnLeuMetLysAspPro 328
1260 TCTTCAGAAATTCCTGATGAATTTATATGTCCTCAATTAATGAGAACTTAAGAAAGATCCG 1319
329 ValIleLeuAspArgGlyTyrSerTyrGlnLysGlnLysMetGlnMetGlnMetTrpIleSerLys 348

1320 GTCATGCATCAGATGCTGATTTTCAATGAAAGAAAGCAATGGAATTTGATCAGCCAA 1379
349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
1380 AAGAAACGTACAGCTCCATGCAAAATCTTGTCTTCTTCAAGCGGTACTTACACCAAT 1439
369 ArgThrLeuLysMetAlaIleAsnArgTrrPLeuGlnThrHisGlnLys 384
1440 AGGACTCTGAAATGGCCATCATATGATGCTGAGAGACACCAAAAG 1487
RESULT 7
US-10-474-291-40
Sequence 40, Application US/10474291
Publication No. US20040132043A1
GENERAL INFORMATION:
APPLICANT: AZIMZAI, Yalda; NU-YOUNG, Janice K.
APPLICANT: BATRA, Sajeev; BAUGHN, Mariah R.
APPLICANT: BECHA, Shanya D.; BOROMSKY, Mark L.
APPLICANT: BURFORD, Neil; DING, Li
APPLICANT: ELLIOTT, Vicki S.; EMERLING, Brooke M.
APPLICANT: GANDHI, Ameena R.; GRETZEN, Kimberly J.
APPLICANT: GRIFFIN, Jennifer A.; HARALIA, April J.A.
APPLICANT: HONCHELL, Cynthia D.; LAU, Preeti G.
APPLICANT: LEE, Soo Yeun; LU, Dyung Alina M.
APPLICANT: ARVIZU, Chandra S.; RAMKUMAR, Jayalaxmi
APPLICANT: REDDY, Roopa M.; SANJUNWALA, Madhusudan M.
APPLICANT: TANG, Y. Tom; CHAWLA, Narinder K.
APPLICANT: WANG, Yu-Mei E.; WARREN, Bridget A.
APPLICANT: XU, Yuming; YANG, Junning
APPLICANT: YAO, Monique G.; YUE, Henry
APPLICANT: ZEBARJADIAN, Yeganeh
TITLE OF INVENTION: PROTEINS ASSOCIATED WITH CELL GROWTH, DIFFERENTIATION, AND DEATH
FILE REFERENCE: PI-0417 USN
CURRENT APPLICATION NUMBER: US/10/474,291
CURRENT FILING DATE: 2003-10-06
PRIOR APPLICATION NUMBER: PCT/US02/11152
PRIOR FILING DATE: 2002-04-05
PRIOR APPLICATION NUMBER: US 60/349,705
PRIOR FILING DATE: 2002-01-15
PRIOR APPLICATION NUMBER: US 60/295,263
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: US 60/295,340
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: US 60/293,727
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: US 60/291,846
PRIOR FILING DATE: 2001-05-18
PRIOR APPLICATION NUMBER: US 60/291,662
PRIOR FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: US 60/287,228
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/286,820
PRIOR FILING DATE: 2001-04-26
PRIOR APPLICATION NUMBER: US 60/283,294
PRIOR FILING DATE: 2001-04-11
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PERL Program
SEQ ID NO 40
LENGTH: 1773
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No: 7483131CB1
US-10-474-291-40
Alignment Scores:
Pred. No.: 7.8e-225 Length: 1773
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2

QY 81 TrpAnthrgluanglyglmetleu1aValmetgluglnProserglySerProVal 100
Db 390 TGGAAATCTGAAAATGACAGATGCTGGCACTGATGAAACGCTTAAGTGGACGCCCTGTG 449
QY 101 ArgValCysglInPheSerProAspSerThrCysleu1aSerglyVala1aaspGlyThr 120
Db 450 AGGGTTGGCCAGTTTCCCGACTCCACGTTGTCATGAGGGGCGAGCTGATGGAAC 509
QY 121 ValValleutrpAna1aaglInserThrlyleu1ytrargCysglYserVallyAspGly 140
Db 510 GTGGTTTGTGGATGACAGCTCATACAAATTAATAGATGTGATGTTAAAGATGGC 569
QY 141 Serleu1a1aCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCys 160
Db 570 TCCTTGGGCGATGCTGCTTTCTCTTAATGAAAGCTTTTGTGACGCTGCTCATGT 629
QY 161 GlyAspLeuThrVal1TPAspAspLysMetArgCysleu1aSerglyVala1aasp 180
Db 630 GGTGATTTAACAGTGTGGATGATATAAGTGTCTGATGATGAAAAAGCACATGAT 689
QY 181 LeuGly1lethrCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
Db 690 CTTGGAATTAACCTCTGCGATTTTCTTCAACAGCAGTTTCTGATGAGAAACAAGTCTT 749
QY 201 GlnPhePheArgleu1a1aSerCysGlyglInAspCysglInVallyle1lethrleValSer 220
Db 750 CAGTTTTCAGCTGGCATCATGTGTGAGATGTCCAAAGTTCAAATTTGGATTGTTCT 809
QY 221 PheThrHis1leLeu----- 225
Db 810 TTTAACCATATCTTAGTTTGAATTAATAATATAAGTACAGTAGTGGGCACCTGTCT 869
QY 225 ----- 225
Db 870 CCTGTTCGCTGTGCTTTTCCCATGATGGCAGATGTAAGTCTCAGGCTCAGTGAT 929
QY 225 ----- 225
Db 930 AAGTCTGCTAGTATATGATATGATATGAGAAATATATCTTCAACATGATGACAGC 989
QY 225 ----- 225
Db 990 ACCAGTATGTCACAACTTGTGCTTTGACCTTAATACCTTTTACTGTACTGTCTCA 1049
QY 226 -----AlaArgArg 228
Db 1050 ATGCAAAAACAGTGAACATCTGGCAATTTGACCTGAACACCTTTGGCAAGCAAGAGC 1109
QY 229 ThrGluHsglnLeuLysGlnPheThrGluAspTrpSerGluGluVala1aSerThrTrp 248
Db 1110 ACAGAACCTCAGCTGAGAACAAATTAACGAAATTTGTGTGAGAGAGATGTTCAACATGG 1169
QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGly1lePheLysMetAsnAsn1laaspGly 268
Db 1170 CTTTGTGACAAAGATTTTAAAGATCTTGTGTATTTTCAAGATGAATTAACATTTGATGCA 1229
QY 269 LysglInleuLeuAsnLeuThrLysglInserLeu1aaspAspLeuLys1leGlnserLeu 288
Db 1230 AAAAGACGTTGATCTTACAAAAGAAAGTCTGGCTGATGATTTGAAAATTTGAATCTCTA 1289
QY 289 GlyLeuAspSerLysValleuArgly1leGluGlnleuAspGlnLysVallySerleu 308
Db 1290 GGACTGCTAGTAAAGTCTGAGGAAATTTGAAAGCTCAGAGCCAAAGTTAAATCCCTT 1349
QY 309 SerSerGly1leProAspGluPhe1leCysPro1leThrArgGluLeuMetLysAspPro 328
Db 1350 TCTTCAGAAATTCCTGATGAATTTATATGTCGAATTAACCTAGAACTTATGAAGATCCG 1409
QY 329 Val1lealaserAspGlyTriserThrGluLysglIna1aMetGluAsnTrp1leSerLys 348
Db 1410 GTCAATCGCATCAGATGGCTATTCATATGAAAAAGAAAGCAAGAAAAATTTGGATGACAAA 1469

QY 349 LysLysArgThrSerProMetThrAsnLeuValleuProSerAlaValleuThrProAsn 368
Db 1470 AAGAAAGTACAAAGTCCCATGACAAATCTTGTCTTCTTCCACGCGTATTAACCAAA 1529
QY 369 ArgThrLeuLysMetAla1leAsnArgTrpLeuGluThrHisglInLys 384
Db 1530 AGGACTCTGAAAATGGCCATCAATATGATGCTGGAGACACCAAAAG 1577

RESULT 9
US-09-971-392-207
; Sequence 207, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Pearson, Cecelia I.
; APPLICANT: Cocks, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; CURRENT FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 207
; LENGTH: 1817
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Template ID: 158923.9
US-09-971-392-207

Alignment Scores:
Pred. No.: 8,11e-225 Length: 1817
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: Gaps: 1

US-10-077-111-13 (1-384) x US-09-971-392-207 (1-1817)

QY 1 MetVallyserleu1eh1sthrleu1aAspHisGlyAspAspValaAspCysAlaPhe 20
Db 148 ATGGTAAGTCACTGATCAACATTAAGTATGATGATGATGATGATGATGATGATGATG 207
QY 21 SerPheSerleuLeu1aThrCysSerleuAspLysThr1leArgLeuThrYrSerleuArg 40
Db 208 TCCTTTCCCTCTTGGCTACTGCTCTTGGACAAACATTCGCTGATCTGTACGT 267
QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyra1aValHisCysCys 60
Db 268 GACTTATCTGAACTGACCAATTCATTAAGTTTCAATGATGATGATGATGATGATGATG 327
QY 61 CysPheSerProSerGlyHis1leLeu1aSerCysSerThrAspGlyThrValleu 80
Db 328 TGTTTCCCTCCCTGACAGCAATATTTGGCATCGGTTCACAGATGGTATCACATGCTCTA 387
QY 81 TrpAnthrgluanglyglmetleu1aValmetgluglnProserglySerProVal 100
Db 388 TGGAAATCTGAAAATGACAGATGCTGGCACTGATGAAACGCTTAAGTGGACGCCCTGTG 447
QY 101 ArgValCysglInPheSerProAspSerThrCysleu1aSerglyVala1aaspGlyThr 120
Db 448 AGGGTTGGCCAGTTTCCCGACTCCACGTTTGGCATCAAGGGGCGAGCTGATGGAAC 507
QY 121 ValValleutrpAna1aaglInserThrlyleu1ytrargCysglYserVallyAspGly 140
Db 508 GTGGTTTGTGGATGACAGCTCATACAAATTAATAGATGTGATGTTAAAGATGGC 567
QY 141 Serleu1a1aCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCys 160

```

Db      568 TCCTTGCGGCGATGTCATTTTCTCTTAATGAAAGCTTCTTGTGCACTGCTCTCATGT 627
Qy      161 G1yAaPLeuThrVal1TTPaAPaPlyMeArGySeLeuH1sSerG1uYalAhiAaP 180
Db      628 GGTGATTTAACTGATGGATGATTAATAAGAGTGTCTGATATGTGAAAAAGACATGAT 667
Qy      181 LeuG1Y1LeThrCySeCyAaAPhSeSeSerG1nProValSeArSpG1yG1uG1nG1yLeu 200
Db      688 CTTGGAAATTACCTGCGCGATTTTCTTCTCAACGCTTCTGATGGAGAAAGAGCTCT 747
Qy      201 G1nPhAPhArGLeuA1aSeCySeG1yG1nAaSPCySeG1nVal1yS1eTTP11eVal1Ser 220
Db      748 CAGTTTTCGACTGCGCATATGTGTCAGAGATTGCCAAGTCAAAATTTGCAATTTGTTCT 807
Qy      221 PheThH1s11eLeu----- 225
Db      808 TTTACCATATCTTAGGTTTGAATTAAATATAAAGTACACGTAGTGGCACTGTGCT 867
Qy      225 ----- 225
Db      868 CCTGTTCTGGCTTGTGCTTTTCCATGATGGCAGATGCTAGTCTCAGGGTCAGTGAT 927
Qy      225 ----- 225
Db      928 AAGTCTGTCACTAGTATATGATTAATACTAATATCTGAGAATATTAATTCTACACATTTGACTCAGCAC 967
Qy      225 ----- 225
Db      988 ACCAGTATGTCAACAATGCTGTGCTTTTGACACTAATACCTTTTAACTTGCTACTGATGTTCA 1047
Qy      226 -----11AATGATG 228
Db      1048 ATGACAAACAACTGTAACATCTGGCAATTTGACCTGGAAACACTTGGCCAAAGAGAGC 1107
Qy      229 ThRG1uH1sG1nLeuYsG1nPhThRG1uAaPTrSeRG1uG1nVal1aSeThTrTP 248
Db      1108 ACGAAACATCACTGTAAGCAATTTACCGAAGATTTGTCAGAGAGAGATGCTCAACATGG 1167
Qy      249 LeuCyA1aG1nAaPLeuYsAaPLeuVal1G1Y11ePheYsMeArAaAaA1eAaPGLY 268
Db      1168 CTTGTGCAACAAGATTTAAAGATCTTGTGTATTTCAAGATGAAATACATTTGATGGA 1227
Qy      269 LysG1uLeuLeuAaAaLeuThrY1yG1uSeLeuA1aAaPLeuYs11eG1uSeLeu 288
Db      1228 AAAGAACGTGTGATCTTCAAAAAGAAAGCTGCTGATATTTGAAAATTTGATCTCTTA 1287
Qy      289 G1yLeuArSeSerYsVal1euaRgYs11eG1uG1nLeuArGThrYsVal1ySeSerLeu 308
Db      1288 GGACTGCGTAGTAAGTGGCTGAGAAATTTGAAAGCTCAGAACCAAGTTAAATCCCTT 1347
Qy      309 SerSeRG1Y1eProAaPGLuPhel1eCySPro11eThArG1uLeuMeC1yAaPPro 328
Db      1348 TCTTCAGAAATTCCTGATGAATTTATATGTCCAATACCTAGAGAACTTAGAAAGATCCG 1407
Qy      329 Val11eAlaSeArSpG1Y1YsSerY1yG1uYsG1uA1aMeC1uAaA1eTTP11eSerYs 348
Db      1408 GTATGCAACAAATGCTATTCATATGAAAAGAAAGCAATGAAAATTTGATCACACAA 1467
Qy      349 LysLysArGThrSeSerProMetThraAaLeuVal1eProSeR1aVal1eUThrProAa 368
Db      1468 AAGAAACGTACAAATCCCATGACAAATCTTGTCTTCTTCAAGGATCTTACCAACAA 1527
Qy      369 ArgThrLeuYsMeAr1a1eAaAaRGrPLeuG1uThrH1sG1nYs 384
Db      1528 AGGACCTGAAAATGGCCATCAATAGATGGCTGAGAGCACACCAAAAG 1575

RESULT 10
US-10-037-270-768
; Sequence 768, Application US/10037270
; Publication No. US20030104529A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
```

```

; APPLICANT: Aaundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Mehman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yundong
; APPLICANT: Wang, Dunru
; APPLICANT: Wang, Zhwei
; APPLICANT: Tillinghast, John
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
; FILE REFERENCE: 784C1P2B
; CURRENT APPLICATION NUMBER: US/10/037,270
; CURRENT FILING DATE: 2002-01-04
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 768
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164)..(1594)
US-10-037-270-768

Alignment Scores:
Pred. No.: 8.3e-225 Length: 1844
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
Gaps: 1

US-10-077-111-13 (1-384) x US-10-037-270-768 (1-1844)
Qy      1 MetVal1ySeLeu11eThrThrLeuA1aAaP1sG1yAaAPVal1aAaCySAlaPhe 20
Db      164 ATGTGAACCTGATTTTACACATTTAGCTGATCAATGTCAGATGTCAATGCTGTGCTTC 223
Qy      21 SerPheSeLeuLeuA1aThrCySeSerLeuAaP1yThr11eArGLeuYsSerLeuArG 40
Db      224 TCTTTTCCCTCTTGCTGCTACTTGTCTTGGCAAAACAAATTCGCTGACTGTTAGCT 283
Qy      41 AaPPhThRG1uLeuProH1sSerProLeuYsPheH1eThrY1aVal1aHiSeCyS 60
Db      284 GACTTATCTGAACCTGCCACATTTCTCATTTGAAGTTTCAATCACTATGCTGTCCACTGCTTC 343
Qy      61 CySPheSeProSeRG1Y1s11eLeuA1aSeCySeSerThAaSPG1YThrThrVal1e 80
Db      344 TGTTTTCCCTTCCAGACATATTTTGGCATGTCTTCAACAGATGTTCCACTGTCTCTTA 403
Qy      81 TrPaenThRG1uAaNG1yG1MeC1euaA1aMeC1uG1nProSeRG1YSeProVal 100
Db      404 TGAATATCTGAAGAAATGGAACAGATGTGGCAGATGGAACACACTAGTGGAGCCCTGG 463
Qy      101 ArgValCySeG1nPhSeSerProAaPSeThrCySeLeuA1aSeRG1YAla1aAaPGLYThr 120
Db      464 AGGATTTGCCAGATTTTCCCAAGACTCCAGCTGTGTGGCATCAGGGGCAAGCTGATGAACT 523
Qy      121 Val1a11eUThrAaA1eG1nSeSerY1yR1aG1uYsR1aG1ySeVal1yAaPGLY 140
Db      524 GTGTTTGTGGAATGCAACATCATCAAAATTAATATAGATGTGTGATTAAGATGGC 583
```

QY 141 SerLeuAlaIaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCys 160
| | | | |
Db 584 TCCTGGCGGCGATGTCATTTCTCTAATGGAAGCTTCTTGTCACGTGCTCTCAATGT 643
QY 161 GlyAspLeuThrValITTPAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
| | | | |
Db 644 GGTGATTTAAACAGTGTGGATGATAAATAGAGGTGTCTGATGTGAAAAAGCACATGAT 703
QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlyLeu 200
| | | | |
Db 704 CTTCGAATTAACCTGCTGCGATTTCTTTCACAGCCAGTTCTGTATGGAGAACAGGCTTT 763
QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrrPileValSer 220
| | | | |
Db 764 CAGTTTTTTCGACCTGGGATCATGTGTGAGGATTTGCCAAGTCAAAATTTGGATTGTTCT 823
QY 221 PheThrHisIleLeu----- 225
| | | | |
Db 824 TTTACCCATATCTTAGGTTTGAATTAATAATATAAGTACATGAGTGGCGACTGTGCT 883
QY 225 ----- 225
| | | | |
Db 884 CTTGTTCTGGCTGTGCTTTTCCCATGAGTGGCAGATGCTAGTCTCAGGGTCACTGAGAT 943
QY 225 ----- 225
| | | | |
Db 944 AAGTCTGTCAATAGTATATGATATCTAATACGAGAAATATCTTCAACATTTGACTCAGCAC 1003
QY 225 ----- 225
| | | | |
Db 1004 ACCAGTATGTCAACAATGTCCTTTTGCACTAATACCTTTACTTGTACTGTTCA 1063
QY 226 -----AlaArgArg 228
| | | | |
Db 1064 ATGAGCAAAACAGTGAAATCTGSCAATTTGACCTGAAACACTTGGCCAGAGAGAC 1123
QY 229 ThrGlnIleGlnLeuLysGlnPheThrGlnAspTrrSerGlnGlnValIleSerThrTrr 248
| | | | |
Db 1124 ACAGAAATCATCGCTGAAAGCATTTTACCAAGATTTGCTCAGAGAGAGATGTCTCAACATGG 1183
QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
| | | | |
Db 1184 CTTGTGTCAACAAAGTTTAAAGATCTTGTGTATTTTCAAGATGAATTAACATTTGATGGA 1243
QY 269 LysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGlnSerLeu 288
| | | | |
Db 1244 AAAAGAACTGTTGAATCTTAAAGAAAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTA 1303
QY 289 GlyLeuArgSerLysValLysArgLysIleGlnGlnLeuArgThrLysValLysSerLeu 308
| | | | |
Db 1304 GGAATGGGTAGTAAAGTCTAGAGAAATTTAGAGCTCAGAGACCAAGTTAAATCCCTT 1363
QY 309 SerSerGlyIleProAspGlnPheIleCysProIleThrArgGlnLeuMetLysAspPro 328
| | | | |
Db 1364 TCTTCAGGAATTCCTGATGAATTTATATGTCAAATACCTAGAGAACTTATGAAAGATCCG 1423
QY 329 ValIleAlaSerArgGlyTrrSerTrrGlnLysGlnLysAlaMetGlnAsnTrrPileSerLys 348
| | | | |
Db 1424 GTCTATCCATCATAGTGTATTTCAATGAAAGAAAGCAATGAAATTTGATCAGCAAA 1483
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValIleLeuThrProAsn 368
| | | | |
Db 1484 AAGAAAGTACAGATCCCATACAAATCTTCTTCTTCAAGGCGTACCTTACCAAAAT 1543
QY 369 ArgThrLeuLysMetAlaIleAsnArgTrrLeuGlnThrHisGlnLys 384
| | | | |
Db 1544 AGGACTCTGAAATAGCCATCATATGATGCTGAGAGACACCAAAAG 1591
RESULT 11
US-10-117-722-768
; Sequence 768; Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Zhang, Jie
APPLICANT: Dmanac, Radoje T.
TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
FILE REFERENCE: 784CIP2BCIP
CURRENT APPLICATION NUMBER: Polypeptides
CURRENT FILING DATE: 2002-04-04
PRIOR APPLICATION NUMBER: US/10/117,722
PRIOR FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1104
SOFTWARE: pc_fl_genes Version 1.0
SEQ ID NO 768
LENGTH: 1844
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (164)..(1594)
US-10-117-722-768
Alignment Scores:
Pred. No.: 8.3e-225 Length: 1844
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 17 Gaps: 1
US-10-077-111-13 (1-384) x US-10-117-722-768 (1-1844)
QY 1 MetValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValaAsnCysCysAlaPhe 20
| | | | |
Db 164 ATGTGAAATCGATTCACACATTAAGCTGATCATGTGACATGTCATGCTGTGCTTTC 223
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTrrSerLeuArg 40
| | | | |
Db 224 TCCTTTCCCTCTTGGTACTTGTCTCTTGGACAAACAAATTTGGCTGATCTGTTACGT 283
QY 41 AspPheThrGlnLeuProHisSerProLeuLysPheHisThrTrrAlaValHisCysCys 60
| | | | |
Db 284 GACTTACTGAACTGCCACATCTCTCCATGAAATTCATACCTATGCTGTCCACTGCTGC 343
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrTrpValLeu 80
| | | | |
Db 344 TGTTCCTCCCTTCAGACATATTTTGGCATTCGTTGCAACAAATGGTATCCACTGTCTTA 403
QY 81 TrpAsnThrGlnAsnGlyGlnMetLeuAlaValMetGlnGlnProSerGlySerProVal 100
| | | | |
Db 404 TCGAATTACTGAAAATGACAGATCTCTGCGATGATGAAACAGCTTACTGCGACCCCTGTG 463
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValaAlaAspGlyThr 120
| | | | |
Db 464 AGGGTTTGCAGATTTTCCCGAGACTCCACGTGTTGGCATCAGGGGCGAGCTGATGAACT 523
QY 121 ValValLeuLysTrpAsnAlaGlnSerTrrLysLeuTrrArgCysGlySerValLysAspGly 140
| | | | |
Db 524 GTGGTTTGTGGAATGACAGTCAATACAAATTAATATGATGTGTGTTAAAGATGGC 583
QY 141 SerLeuAlaIaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
| | | | |
Db 584 TCCTTGGCGGCGATGTGATTTTCTCTTAATGGAAGCTTCTTGTGCTGCTCTCAATGT 643
QY 161 GlyAspLeuThrValITTPAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
| | | | |
Db 644 GGTGATTTAAACAGTGTGGATGATTAATAATAGAGGTGTCTGATGTGAAAAAGCACATGAT 703

```
QY 181 LeuGlyIleThrcyCyaaPheSerSerGlnProValSerApGlyGluGlnGlyLeu 200
| | | | |
DB 704 CTGGAAATTAACCTGCTGCGATTTTCTTTCACAGCCAGTTCTGATGGAACAAGGCTT 763
QY 201 GlnPhePheArgLeuAlaSerCyseGlyGlnAapCyseGlnValIysIleTPrIleValSer 220
| | | | |
DB 764 CAGTTTTCGACTGGCATATGTGTGACGATTTGCCAATTCGCAAAATTTGGATTGTTCT 823
QY 221 PheThrHisIleLeu----- 225
| | | | |
DB 824 TTTAACCATATCTTAGGTTTGAATTAAATATATAAGTACAGTGGGCACTGTGCT 883
QY 225 ----- 225
DB 884 CCTGTTCTGCTTGTGCTTTTCCATGATGGCAGATGCTCAAGGTCAGTGAT 943
QY 225 ----- 225
DB 944 AAGTCTGTCTAATGATATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1003
QY 225 ----- 225
DB 1004 ACCAGGATGTCAACAATTGCTTTTGACCTAATACCTTTTACTTGTCTAGTGTCA 1063
QY 226 -----Aaa-gArg 228
| | | | |
DB 1064 ATGGAACAAAACAGTGAACATCTGCAATTGACCTGGMAACCTTGGCCAAAGAGGC 1123
QY 229 ThrGluHISGlnLeuIysGlnPheThrGluAapTrpSerGlyGluValIleSerThrTrp 248
| | | | |
DB 1124 ACAGAACATACGCTGAAGCAATTTACCAAGATTTGGCAGAGAGAGATGCTCAACATGG 1183
QY 249 LeuCyseAlaGlnAapLeuIysAapLeuValGlyIlePheIysMetAanIleAapGly 268
| | | | |
DB 1184 CTTTGTCACAAAGATTTAAAGATCTTGTGTATTTTCAAGATGATTAACATTGATGGA 1243
QY 269 LysGluLeuLeuAapLeuThrIlyGluSerLeuAlaAapAapLeuIysIleGluSerLeu 288
| | | | |
DB 1244 AAAGAACTGTGTGAATCTTCAAAAAGAAAGTCTGCTGATTTGAAAATTTGAATCTCTA 1303
QY 289 GlyLeuAapSerIysValIleuArgIysIleGluGluLeuAapThrIlyIysSerLeu 308
| | | | |
DB 1304 GGACTGTGATGAAAGGCTGAGGAAATGAAAGGCTCAGACCAAGGTTAAATCCCTT 1363
QY 309 SerSerGlyIlePhePheGlnPheIleCyseProIleThrArgGluLeuMetIysAapPro 328
| | | | |
DB 1364 TCTTCAGGAATTCCTGATGATTTATATGTCATATAGTGAAGAATTAAGAAAGATCCG 1423
QY 329 ValIleAlaSerAapGlyIlySerIlyGluIysGluAlaMetGluAanTPrIleSerIys 348
| | | | |
DB 1424 GTCATCGCATCAATGGCTATTTCAATATGAAAGGAAAGCAATGGAATTTGATCAGCAAA 1483
QY 349 LysIysAapThrSerPheMetThrAanLeuValLeuProSerAlaValIleuThrProAan 368
| | | | |
DB 1484 AAAGAAAGTACAGTCCATGACCAATCTTGTCTTCTTCAGGGGATTAACCAAAAT 1543
QY 369 ArgThrLeuIysMetAlaIleAanArgTrpLeuGluThrHisGlnIys 384
| | | | |
DB 1544 AGGACTCTGAAAATGGCCATCAATAGATGGCTGAGAGACACCAAAAG 1591
RESULT 12
US-10-077-111-14
; Sequence 14, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jili
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
```

```
QY 1 PRIOR FILING DATE: 2001-05-29
| | | | |
DB 136 ATGGTGAACCTGATTCACATTAAGCTGATGCTGAGAGATGTCACATGCTGTGCTTC 195
QY 21 SerPheSerLeuAapLeuAlaThrCyseSerLeuAapIlyThrIleArgLeuIlySerLeuArg 40
| | | | |
DB 196 TCTTTTCCCTCTTGGCTACTGCTGCTCTTGGACAAAACAATTCGCTGACTGTTACGT 255
QY 41 AapPheThrGluLeuProHisSerProLeuIysPheHisThrIlyAlaValHisCyseCyse 60
| | | | |
DB 256 GACTTACTGAAGCTGCAATTCCTCATTTGAAGTTTCATACCTATGCTGTCACTGCTG 315
QY 61 CysePheSerProSerGlyIlyIleLeuAlaSerCyseSerThrAapGlyIlyThrValIleu 80
| | | | |
DB 316 TGTTCCTCCCTTCAGAGACATATTTGGCATGTGTTCACAGAGTGTACACTGTGCTTA 375
QY 81 TrpAanThrGluAanGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
| | | | |
DB 376 TGGAAATCTGAAAATGAGACAGATGCTGCAAGTATGAGAACACCTATGTCAGCCCTG 435
QY 436 AGGATTTGCCAGTTTCCCAAGCTCCAGTGTGTCATCAGGGGCACTGATGGAAT 495
QY 121 ValValLeuThrAanAlaGlnSerIlyIysLeuIlyArgCyseGlySerValIysAapGly 140
| | | | |
DB 496 GTGGTTTGTGAAATGACAGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 555
QY 141 SerLeuAlaIleCyseAlaPheSerProAanGlySerPhePheValThrIlySerSerCyse 160
| | | | |
DB 556 TCTTGGCGGCAATGCTCATTTTCTCTTAATGAAAGCTTTTGTGCTAGCTGCTCTATGT 615
QY 161 GlyAapLeuThrValTrpAapAapIysMetArgCyseLeuHisSerGlyIlyValHisAap 180
| | | | |
DB 616 GGTGATTTAAACGTGTGGATGATTAATGAGATGCTCATATGTCATGAAACAAAGTCTT 675
QY 181 LeuGlyIleThrcyCyaaPheSerSerGlnProValSerApGlyGluGlnGlyLeu 200
| | | | |
DB 676 CTGGAAATTAACCTGCTGCGATTTTCTTTCACAGCCAGTTCTGATGGAACAAGGCTCT 735
QY 201 GlnPhePheArgLeuAlaSerCyseGlyGlnAapCyseGlnValIysIleTPrIleValSer 220
| | | | |
DB 736 CAGTTTTCGACTGGCATATGTGTGACGATTTGCCAATTCGCAAAATTTGGATTGTTCT 795
QY 221 PheThrHisIleLeu----- 225
| | | | |
DB 796 TTTAACCATATCTTAGGTTTGAATTAAATATATAAGTACAGTGGGCACTGTGCT 855
QY 225 ----- 225
US-10-077-111-13 (1-384) x US-10-077-111-14 (1-1908)
```

```

Db      856 CCTGTTGCTTGTGCTTTTCCGANGATGGCAGATGTAAGTCTCAGGTCAGTGAT 915
Qy      225 -----
Db      916 AAGTCTGCATAGTATGATAGTAATACTGAGAAATACTTTCACACATTGACTCAGCAC 975
Qy      225 -----
Db      976 ACCAGATATGTCACAACTTGCTTTGGCACCTAATACCCTTTACTTGTAAGTTCA 1035
Qy      226 -----
Db      1036 ATGACAAACAGTGAACATCTGCAATTTGACCTGAAACATTTGCCAAGCAGAGGCG 1095
Qy      229 ThrluhlsleuLysGlnPheThrGluAspTrpSerGluGluValIleSerThrTrp 248
Db      1096 ACAGAACTCAGCTGAGACATTTTACGAAATTTGCTAGAGAGATGCTTCACACATGG 1155
Qy      249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
Db      1156 CTTTGTCACAAAGATTAAAGATCTTGTAATTTTCAAGATGAATTAACATTGATGGA 1215
Qy      269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIle----- 285
Db      1216 AAGAACTGTTGAATCTTACAAAGAAAGTCTGGCTGATGATTTGAAATTTGGCTGGAGT 1275
Qy      285 -----
Db      1276 CCTTGGCAGTGTGCTGCTCAGCTGACAGCTTCAACTCTGGGCTCAAGTATCCTCTTA 1335
Qy      286 -----GluSerLeuGlyLeuArgSerLysValLeuArgLysIleGluGluLeuArg 302
Db      1336 CCTGGGCTCATCTCTAGAGCTGCGTAGTAAGTCTGAGAGAAATTTGAAGACTCAGG 1395
Qy      303 ThrLysValLysSerLeuSerSerGlyIleProAspGluPheLysCysProIleThrArg 322
Db      1396 ACCAAGGTTAAATCCCTTTCTTTCAGGAATCTGATGAATTTATATGCTCCAAATACTAGA 1455
Qy      323 GluLeuMetLysAspProValIleAlaSerAspGlyTrpSerTrpGluLysGluAlaMet 342
Db      1456 GAACCTTGAAGAATCCGCTGATGCGACATGCGCTAATTAATGAAAGAAAGCAATG 1515
Qy      343 GluAsnTrpLysSerLysLysValArgThrSerProMetThrAsnLeuValLeuProSer 362
Db      1516 GAAATTTGATCAGCAAAAAGAAAGCTACAGTCCCATGACAAATCTTTGCTCTTCA 1575
Qy      363 AlaValLeuThrProAsnArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHis 382
Db      1576 GCGGTACTTACACCAAAATAGGACTCTGAAATGGCCATCATATGAGCTGGAGACACAC 1635
Qy      383 GlnLys 384
Db      1636 CAAAG 1641

```

```

RESULT 13
US-10-077-111-6
; Sequence 6, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jili
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-414US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; PRIOR FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: 60/269,366
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6

```

```

; LENGTH: 1901
; TYPE: DNA
; ORGANISM: MOUSE
US-10-077-111-6

Alignment Scores:
Pred. No.: 1,07e-181 Length: 1901
Score: 1616.50 Matches: 312
Percent Similarity: 72.06% Conservative: 31
Best Local Similarity: 65.55% Mismatches: 40
Query Match: 78.97% Indels: 93
DB: 13 Gaps: 3

US-10-077-111-13 (1-384) x US-10-077-111-6 (1-1901)
Qy      1 MetValLysLeuIleHsThrLeuAlaAspHisGlyAspAspValaAsnLysCysAlaPhe 20
Db      19 ATGTTGAGTTGATTCACAGCTGCTGATGATCAGGAGATGACCTCAGCTCTGCGCTTC 78
Qy      21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTrpSerLeuArg 40
Db      79 TCGGCTGCTCTCTGCGCACCTGCTCTTGGACAGACCATCCGCTGTACTCTCCCTAAGT 138
Qy      41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyraAlaValHisCysCys 60
Db      139 GACTTGTGAACTGCCGTACTCCCGCTGAAGTTCCACACCTATGCTGTCACTGCTGC 198
Qy      61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
Db      199 TGTTCACCTCAGACAGCAAGTTTATGATGCTGCTGACAGACGAGCCAGCCAGCTGCTG 258
Qy      81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db      259 TGGAGCTGCAACGCGACACACACCTGACCGTGTGGAGAGCCGGGTGGAGCCCTGTG 318
Qy      101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
Db      319 CGGCTGCTGCTTTTCCAGACTGCTGCTTACTGACGTCAGGCGCTGCCATGATCC 378
Qy      121 ValValLeuTrpAsnAlaGlnSerTrpLysLeuTrpArgCysGlySerValLysAspGly 140
Db      379 ATTGCTTTGGAAATGACACAGACATACAACTATATGAGTGTGTGTCAGAGATGC 438
Qy      141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCys 160
Db      439 TCATGTGCTGCTGTGCTTTTCCGATGAGGCTCTTTGCACTGCTCTCGGCGC 498
Qy      161 GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
Db      499 GGGACCTTGACAGTGGAGTGAACAGATGAGGTCTTACACAGCGAAGGCGCACGAT 558
Qy      181 LeuGlyLysThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200
Db      559 CTCGGATCAGCTGCTGCTGCTTTCTCTCAGAGCTCTCTCGGCGGAGAA---GGCTC 615
Qy      201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer 220
Db      616 CAGCTTACCAAGTTGGGTGATGTGTCAGAGCTGGAATCAAACTGGGCTGTTACT 675
Qy      221 PheThr----- 222
Db      676 ATTAACCGGTCTTGAAGCTTTGAATTAATATTAAGACACACTAAGTGGGCACTGGCC 735
Qy      222 ----- 222
Db      736 CCTGTTGCGCTGTGCTTTTTCACATGATGAAAGATGCTTGCAATCGGGGTAGTGAT 795
Qy      223 -----HisIleLeuAlaArgArg 228
Db      796 AAATCTGTCATCATACATGATGCGCCCTCAGAGTGTGCTACACAGCTGACTCAGCAT 855
Qy      229 ThrGluHis----- 231

```

| | | | |
|----|------|---|------|
| Db | 856 | AACAGAGTATGTTACGACTTTGGCCGTTTGACACCCAAACACTCTCTTAAGTCTACTGCTTCA | 915 |
| OY | 231 | ----- | 231 |
| Db | 916 | ATGCAACAAGACAGTACATTTGGCAGTTTGACTGCAAAACCTTGCACAGCAAGAAC | 975 |
| OY | 232 | -----GlnLeuLysGlnPheThrGluAspTrpSerGluGluValValSerThrTrp | 248 |
| Db | 976 | ATGAACGACCCGCTGAAAACATTTCACTAGAAATAGTTCAGAGAGAGATGTCTCCGTGG | 1035 |
| OY | 249 | LeuCySAIaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly | 268 |
| Db | 1036 | CTTCGCTGCTCAAGCGTTTGGAAAGCTTCGTGGTATTTCAGGGCAAAACAACTGCATGG | 1095 |
| OY | 269 | LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuValGluSerLeu | 288 |
| Db | 1096 | AAAGAACATTTGCATCTACAAAGAAAGCTCTGCGTGTGATTTGAAATGCAATCTCTA | 1155 |
| OY | 289 | GlyLeuAspSerLysValLeuArgLysIleGluGluLeuAspThrLysValLysSerLeu | 308 |
| Db | 1156 | GGGCTGCCACAGCAAAAGTCTCTGAGGAGTATTGAAGAGCTCAAGGGCCAAAGATGATTCCTC | 1215 |
| OY | 309 | SerSerGlyIleProAspGluPheIleCyProIleThrArgGluLeuMetLysAspPro | 328 |
| Db | 1216 | TCTTCGGGAATCTCCGACGAGTCACTCGCCCAACAAACAAAGAACTCATAGAGACCC | 1275 |
| OY | 329 | ValIleAlaSerAspGlyTrpSerYrGluLysGluAlaMetGluAsnTrpIleSerLys | 348 |
| Db | 1276 | GTCATCGCATCAGATGGCTACTCTTCACAGAGAGAAAGCAATGAAAGCTGCATCCACAG | 1335 |
| OY | 349 | LysLysAspThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn | 368 |
| Db | 1336 | AAGAAAGCTACGAGGCCCATGACAAATTTGGCTCTCCCTTCACTGTGACTGACCCCAAC | 1395 |
| OY | 369 | ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrIleGlnLys | 384 |
| Db | 1396 | AGGACACTGAAAGATGGCCATCAACCGATGGCTGGAGACGACAGAGAG 1443 | |

; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5029
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-5029

Alignment Scores:
Pred. No.: 8,33e-65 Length: 446
Score: 627.50 Matches: 122
Percent Similarity: 96.06% Conservative: 0
Best Local Similarity: 96.06% Mismatches: 2
Query Match: 30.65% Indels: 3
DB: 18 Gaps: 1

US-10-077-111-13 (1-384) x US-10-357-930-5029 (1-446)

QY 102 ValCYSGln-PheSerProAspSer--ThrCYsLeuAlaSerGlyAlaAlaAspGlyTh 120
|||
Db 444 GTTGGCCAGTTTTCCTCCAGACTCCCAACGTGTTGGCATCAGGGCAGCTGATGAGAC 385

QY 120 rValValLeuTrpAsnAlaGlnSerTYrLYsLeuTYrArgCYsGlySerValLYsAspG 140
|||
Db 384 TGTGGTTTGTGAAATGCAAGTCATACAAATTATATGATGTGTAGTGTTPAAGATGG 325

QY 140 YSerLeuAlaAlaCYsAlaPheSerProAsnGlySerPhePheValThrgLYsSerSerCY 160
|||
Db 324 CTCCTGGCGGCAGTGTGATTTTCTCCATAGGAAGCTTTTGTCACTGGCTCTCATG 265

QY 160 sGlyAspLeuThrValTrpAspAspLYsMetArgCYsLeuHisSerGlu-LysAlaHisSA 180
|||
Db 264 TGTGATTTAAACAGTGTGGATGATAAATGAGGTGTCTGCATAGTGAATAAAGCACATG 205

QY 180 sPLeuGlyLeThrCYsAspAspPheSerSerGlnProValSerAspGlyGlnGlnGlyL 200
|||
Db 204 ATCTGGAAATTAACCTGCTGCGATTTTCTTCACAGCCAGTTTCTGATGAGAGACAGGTC 145

QY 200 eugInPhePheArgLeuAlaSerCYsGlyGlnAspCYsGlnValLYsIleTrpIleValS 220
|||
Db 144 TTCAGTTTTCGACTGGCATCATGTGTGTCAAGATTGCCAAGTCAAAATTTGGATTGTTT 85

QY 220 erPheThrHisIleLeu 225
|||
Db 84 CTTTACCATATCTTA 68

Search completed: February 5, 2005, 15:08:28
Job time : 719 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 08:35:31 ; Search time 748 Seconds
(without alignments)
9787.839 Million cell updates/sec

Title: US-10-077-111-12
Perfect score: 1272
Sequence: 1 gaattcggttcctaccctgcg.....acaccacaaagtaagaattc 1272

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 3483234

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCF_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCFUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq:*
19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---|
| C 1 | 176.8 | 13.9 | 180 | 9 | US-09-864-761-27687, A Sequence 27687, A |
| C 2 | 132 | 10.4 | 132 | 9 | US-09-864-761-30542 Sequence 30542, A |
| C 3 | 87 | 6.8 | 92 | 15 | US-10-106-698-3381 Sequence 3381, App |
| C 4 | 36 | 2.8 | 36 | 13 | US-10-077-111-19 Sequence 19, App |
| C 5 | 33.6 | 2.6 | 152 | 18 | US-10-425-115-14085 Sequence 14085, A |
| C 6 | 31 | 2.4 | 165 | 17 | US-10-242-535A-49505 Sequence 49505, A |
| C 7 | 31 | 2.4 | 165 | 17 | US-10-085-783A-49505 Sequence 49505, A |
| C 8 | 30.6 | 2.4 | 136 | 9 | US-09-770-696-357 Sequence 357, App |
| C 9 | 29.6 | 2.3 | 136 | 9 | US-09-864-761-27694 Sequence 27694, A |
| C 10 | 29.2 | 2.3 | 177 | 18 | US-10-437-963-57611 Sequence 57611, A |
| C 11 | 29 | 2.3 | 170 | 10 | US-09-960-706-190 Sequence 190, App |

| | | | | | |
|------|------|-----|-----|----|--|
| 12 | 29 | 2.3 | 180 | 15 | US-10-238-075-979 Sequence 979, App |
| 13 | 28.8 | 2.3 | 152 | 9 | US-09-864-761-21108 Sequence 21108, A |
| 14 | 28.8 | 2.3 | 153 | 9 | US-09-864-761-19055 Sequence 19055, A |
| C 15 | 28.6 | 2.2 | 182 | 17 | US-10-242-535A-34626 Sequence 34626, A |
| C 16 | 28.6 | 2.2 | 182 | 17 | US-10-085-783A-34626 Sequence 34626, A |
| C 17 | 28.4 | 2.2 | 60 | 10 | US-09-908-978-12618 Sequence 12618, A |
| C 18 | 28.2 | 2.2 | 182 | 17 | US-10-424-559-64092 Sequence 64092, A |
| C 19 | 28 | 2.2 | 128 | 13 | US-10-027-632-76001 Sequence 76001, A |
| C 20 | 28 | 2.2 | 143 | 17 | US-10-027-632-76001 Sequence 76001, A |
| C 21 | 28 | 2.2 | 143 | 13 | US-10-027-632-76003 Sequence 76003, A |
| C 22 | 28 | 2.2 | 143 | 17 | US-10-027-632-76003 Sequence 76003, A |
| C 23 | 28 | 2.2 | 147 | 10 | US-09-814-355-5092 Sequence 5092, App |
| C 24 | 28 | 2.2 | 147 | 10 | US-09-814-355-11384 Sequence 11384, App |
| C 25 | 28 | 2.2 | 174 | 18 | US-10-674-124A-6680 Sequence 6680, App |
| C 26 | 28 | 2.2 | 177 | 18 | US-10-425-115-23798 Sequence 23798, A |
| C 27 | 28 | 2.2 | 186 | 18 | US-10-437-963-83715 Sequence 83715, A |
| C 28 | 28 | 2.2 | 197 | 10 | US-09-468-147-151 Sequence 151, App |
| C 29 | 28 | 2.2 | 197 | 17 | US-10-319-745-151 Sequence 151, App |
| C 30 | 27.8 | 2.2 | 170 | 9 | US-09-815-242-2740 Sequence 2740, App |
| C 31 | 27.8 | 2.2 | 170 | 9 | US-09-815-242-3443 Sequence 3443, App |
| C 32 | 27.8 | 2.2 | 170 | 17 | US-10-282-122A-5261 Sequence 5261, App |
| C 33 | 27.8 | 2.2 | 170 | 17 | US-10-282-122A-6027 Sequence 6027, App |
| C 34 | 27.8 | 2.2 | 183 | 17 | US-10-433-140A-1 Sequence 1, Appl |
| C 35 | 27.8 | 2.2 | 183 | 17 | US-10-433-140A-3 Sequence 3, Appl |
| C 36 | 27.6 | 2.2 | 183 | 17 | US-10-425-115-31475 Sequence 31475, A |
| C 37 | 27.6 | 2.2 | 184 | 9 | US-09-864-761-20195 Sequence 20195, A |
| C 38 | 27.4 | 2.2 | 146 | 18 | US-10-437-963-92321 Sequence 92321, A |
| C 39 | 27.4 | 2.2 | 192 | 9 | US-09-796-692-83322 Sequence 8322, App |
| C 40 | 27.4 | 2.2 | 192 | 14 | US-10-040-862-83322 Sequence 8322, App |
| C 41 | 27.4 | 2.2 | 192 | 17 | US-10-057-475B-83322 Sequence 8322, App |
| C 42 | 27.4 | 2.2 | 192 | 17 | US-10-154-884B-83322 Sequence 8322, App |
| C 43 | 27.4 | 2.2 | 192 | 18 | US-10-764-324-83322 Sequence 8322, App |
| C 44 | 27.4 | 2.2 | 195 | 17 | US-10-242-535A-2852 Sequence 2852, App |
| C 45 | 27.4 | 2.2 | 195 | 17 | US-10-085-783A-2852 Sequence 2852, App |

ALIGNMENTS

RESULT 1
US-09-864-761-27687/c
; Sequence 27687, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30

```

; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 27687
; LENGTH: 180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC006501.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.72
; OTHER INFORMATION: EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94
; OTHER INFORMATION: NT HIT: U43139.1, EVALUE 1.10e-01
; OTHER INFORMATION: SWISSPROT HIT: Q16760, EVALUE 6.80e-02
US-09-864-761-27687
```

```

Query Match      13.9%; Score 176.8; DB 9; Length 180;
Best Local Similarity 98.9%; Pred. No. 5.8e-41;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 787 CAAGGCGCAGACATCTGAGCAATTTACCGAAGATTGCTGAGAGAGTGTCT 846
    |||||
DB 180 CAAGGCGCAGACATCTGAGCAATTTACCGAAGATTGCTGAGAGAGATGCT 121
    |||||
QY 847 CAACATGGCTTTGTGCAACAAGATTAAAGATCTTGTGTATTTTCAAGATGATPACA 906
    |||||
DB 120 CAACATGGCTTTGTGCAACAAGATTAAAGATCTTGTGTATTTTCAAGATGATPACA 61
    |||||
QY 907 TTGATGAAAAGACATCTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTGAAAATTG 966
    |||||
DB 60 TTGATGAAAAGACATCTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTGAAAATTG 1
    |||||
```

```

RESULT 2
US-09-864-761-30542/c
; Sequence 30542, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
```

```

; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 30542
; LENGTH: 132
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009307.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.76
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.73
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.79
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 0.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.78
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.89
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.75
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 0.68
; OTHER INFORMATION: EST_HUMAN HIT: BE675766.1, EVALUE 2.00e-68
; OTHER INFORMATION: NT HIT: X71133.1, EVALUE 1.20e+00
; OTHER INFORMATION: SWISSPROT HIT: P54860, EVALUE 1.00e+00
US-09-864-761-30542
```

```

Query Match      10.4%; Score 132; DB 9; Length 132;
Best Local Similarity 100.0%; Pred. No. 6.1e-28;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 976 GACTGCGTAGTAAGTCTGAGAAAATGGAAGCTCAGACCAAGTTAATCCCTTT 1035
    |||||
DB 132 GACTGCGTAGTAAGTCTGAGAAAATGGAAGCTCAGACCAAGTTAATCCCTTT 73
    |||||
QY 1036 CTTGAGGAATTCCTGATGATTTATATGTCCAAATACTGAGAACTTATGAAAGATCCGG 1095
    |||||
DB 72 CTTGAGGAATTCCTGATGATTTATATGTCCAAATACTGAGAACTTATGAAAGATCCGG 13
    |||||
QY 1096 TCATGCGCATCAG 1107
    |||||
DB 12 TCATGCGCATCAG 1
    |||||
```

```

RESULT 3
US-10-106-698-3381
; Sequence 3381, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
```

;; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
;; FILE REFERENCE: PA005B1
;; CURRENT APPLICATION NUMBER: US/10/106,698
;; CURRENT FILING DATE: 2002-03-27
;; PRIOR APPLICATION NUMBER: PCT/US00/26524
;; PRIOR FILING DATE: 2000-09-28
;; PRIOR APPLICATION NUMBER: US 60/157,137
;; PRIOR FILING DATE: 1999-09-29
;; PRIOR APPLICATION NUMBER: US 60/163,280
;; PRIOR FILING DATE: 1999-11-03
;; NUMBER OF SEQ ID NOS: 8564
;; SOFTWARE: PatentIn Ver. 3.0
;; SEQ ID NO 3181
;; LENGTH: 92
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc_feature
;; LOCATION: (45)..(45)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (78)..(78)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (86)..(86)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (89)..(89)
;; OTHER INFORMATION: n equals a,t,g, or c
;; NAME/KEY: misc_feature
;; LOCATION: (91)..(91)
;; OTHER INFORMATION: n equals a,t,g, or c
US-10-106-698-3381

Query Match
Best Local Similarity 6.8%; Score 87; DB 15; Length 92;
Best Local Similarity 94.6%; Pred. No. 7.1e-15;
Matches 87; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 591 GGTGATTACAGCTGGGATGATTAAGTGGTCTGCTAGTGAAGGACATGAT 650
DB 1 GGTGATTACAGCTGGGATGATTAAGTGGTCTGCTAGTGAAGGACATGAT 60
QY 651 CTGGATTACCTGCTGCGATTTTCTTCA 682
DB 61 CTGGATTACCTGCTGCGATTTTCTTCA 92

RESULT 4
US-10-077-111-19/c
;; Sequence 19, Application US/10077111
;; Publication No. US20020187492A1
;; GENERAL INFORMATION:
;; APPLICANT: Toddard, C. Gordon
;; APPLICANT: Finger, Joshua N.
;; APPLICANT: Rillema, Jill
;; TITLE OF INVENTION: TBA
;; FILE REFERENCE: 3053-4114US2
;; CURRENT APPLICATION NUMBER: US/10/077,111
;; CURRENT FILING DATE: 2002-02-15
;; PRIOR APPLICATION NUMBER: 60/294,181
;; PRIOR FILING DATE: 2001-05-29
;; PRIOR APPLICATION NUMBER: 60/269,366
;; PRIOR FILING DATE: 2001-02-16
;; NUMBER OF SEQ ID NOS: 25
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 19
;; LENGTH: 36
;; TYPE: DNA
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Primer, JNF
US-10-077-111-19

Query Match
Best Local Similarity 2.8%; Score 36; DB 13; Length 36;
Best Local Similarity 100.0%; Pred. No. 3.3;
Matches 36; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1230 GCCATCAATAGATGCTGGAGACACCAAAAGTAA 1265
DB 36 GCCATCAATAGATGCTGGAGACACCAAAAGTAA 1

RESULT 5
US-10-425-115-140885/c
;; Sequence 140885, Application US/10425115
;; Publication No. US20040214272A1
;; GENERAL INFORMATION:
;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Kovacic, David K.
;; APPLICANT: Zhou, Yihua
;; APPLICANT: Cao, Yongwei
;; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;; TITLE OF INVENTION: Plants
;; FILE REFERENCE: 38-21(53222)B
;; CURRENT APPLICATION NUMBER: US/10/425,115
;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 369326
;; SEQ ID NO 140885
;; LENGTH: 152
;; TYPE: DNA
;; ORGANISM: Zea mays
;; FEATURE:
;; OTHER INFORMATION: Clone ID: MRT4577_5996C.1
US-10-425-115-140885

Query Match
Best Local Similarity 2.6%; Score 33.6; DB 18; Length 152;
Best Local Similarity 56.2%; Pred. No. 39;
Matches 63; Conservative 0; Mismatches 49; Indels 0; Gaps 0;
QY 323 GTGTCAACAGATGATACAGCTGTCTATGGAATGAAATGACAGATGCGAGT 382
DB 148 GTTTCACACAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 89
QY 383 GATGGAACAGCTGATGAGCCCTGTGAGGCTTTGCACTTTTCCACAGAC 434
DB 88 TATGCAATCAGCCAGCGCGGTGACGCGGAGGATTTGCGACGTGCGCCACAGC 37

RESULT 6
US-10-242-535A-49505
;; Sequence 49505, Application US/10242535A
;; Publication No. US20040013663A1
;; GENERAL INFORMATION:
;; APPLICANT: ChondroGene Inc.
;; APPLICANT: Liew, C.C.
;; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
;; FILE REFERENCE: 4231/2005
;; CURRENT APPLICATION NUMBER: US/10/242,535A
;; CURRENT FILING DATE: 2002-09-12
;; PRIOR APPLICATION NUMBER: US 10/085,783
;; PRIOR FILING DATE: 2002-02-28
;; PRIOR APPLICATION NUMBER: US 60/305,340
;; PRIOR FILING DATE: 2001-07-13
;; PRIOR APPLICATION NUMBER: US 60/275,017
;; PRIOR FILING DATE: 2001-03-12
;; PRIOR APPLICATION NUMBER: US 60/271,955
;; PRIOR FILING DATE: 2001-02-28
;; NUMBER OF SEQ ID NOS: 58994
;; SOFTWARE: PatentIn version 3.2
;; SEQ ID NO 49505
;; LENGTH: 165
;; TYPE: DNA
;; ORGANISM: Human
US-10-242-535A-49505

ORGANISM: Escherichia coli
US-10-238-075-979

Query Match
Best Local Similarity 49.7%; Pred. No. 9,7e+02;
Matches 74; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

QY 557 TAATGAAAGCTTTTGTCTGCTGCTTCATGATGATTAACAGTGTGGATGATAA 616
DB 21 TATTGAAACCGGAATGCGCGGTGAGATTTTCGAGGATGAATTCGAAGTGTGATGATGC 80

QY 617 AATGAGGTGTCTCATGATGAAAGACATGATCTTGAATTAACCTGCTGCATTTTC 676
DB 81 GAGCGCTGATGAAATACATGATGACGAAAGCAAAAGATTTTCTTAATCACTGCAATTAATC 140

QY 677 TTCACAGCCAGTTTCTGATGAGAACAG 705
DB 141 ATATCAAGAAATTAAGACGAAAGAGAG 169

RESULT 13
US-09-864-761-21108
Sequence 21108, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
SEQ ID NO 21108

LENGTH: 152
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007721.10
OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.95
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.4
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 0.96
OTHER INFORMATION: EST HUMAN HIT: BE142985.1, EVALU 2.00e-80
OTHER INFORMATION: NT HIT: AB012193.1, EVALU 1.00e-80
OTHER INFORMATION: SWISSPROT HIT: Q13619, EVALU 3.00e-23
US-09-864-761-21108

Query Match
Best Local Similarity 2.3%; Score 28.8; DB 9; Length 152;
Matches 72; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 777 CATATCTTAGCAAGCGCCACAGACATCACTGATGAGCAATTTACCGAAGATTGCTAGAG 836
DB 6 CAGTCCCTGCGCTGTGGCAAGCACGTGTGATTAAGTCCCAAGGAAAGAGTG 65

QY 837 GAGGTGCTCAACAGCGCTTGTGCACAAGATTTAAAGATCTTGTGTAATTTTCAG 896
DB 66 GAAGATGAGAGACAGATTCAATTTTAATGAGAGGTTCAACACAAATTGTTGAATTAAG 125

QY 897 ATGAATACATTGATGAAAGAA 920
DB 126 ATCAATCAATTCAATGAAAGAA 149

RESULT 14
US-09-864-761-19055
Sequence 19055, Application US/09864761
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30


```
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 19055
; LENGTH: 153
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007721.7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 11
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.7
; OTHER INFORMATION: EST HUMAN HIT: BE142985.1, EVALU8 4.00e-81
; OTHER INFORMATION: SWISSPROT HIT: Q13619, EVALU8 8.00e-24
; OTHER INFORMATION: NT HIT: AB012193.1, EVALU8 4.00e-81
US-09-864-761-19055
```

```
Query Match          2.3%; Score 28.8; DB 9; Length 153;
Best Local Similarity 50.0%; Pred. No. 1e+03;
Matches 72; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
```

```
QY 777 CATATCTTAGCAGGCCGACAGAACATCAGCTGAAGCATTTTACGAGATTGTCAGAG 836
    |||||
DB 7 CAGTCCCTGGGCTGTGCAAGACGCTGTGATTAAGTCCCAAGAGAAAGAGAGTG 66

QY 837 GAGTCTCTCAACATGCGCTTTGTGCAAGATTAAAGATCTTGTGATTTTCAAG 896
    |||||
DB 67 GAAGATGAGAGACAGATTCATTTTATGAGAGATTCAAGACAGATTTGTTAGATTAAG 126

QY 897 ATGAATAACATTGATGAGAAAGAA 920
    |||||
DB 127 ATCAATCAATTCAGATGAGAGAA 150
```

RESULT 15

```
US-10-242-535A-34626/c
; Sequence 34626, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: Liew, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242,535A
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34626
; LENGTH: 182
; TYPE: DNA
; ORGANISM: Human
```

US-10-242-535A-34626

```
Query Match          2.2%; Score 28.6; DB 17; Length 182;
Best Local Similarity 51.1%; Pred. No. 1.3e+03;
Matches 67; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
```

```
QY 907 TTGATGMAAAGAACTGTTGAATCTTACAAAGAAAGCTGCTGATGATTTGAAAATTG 966
    |||||
DB 173 TTTATTGCCTCGAAGATGTCAGATCTTACAAATGAAACAAATATATGTTTAAATA 114

QY 967 AATCTTAGACTGCGTAGTAAAGTCTGAGAAAATTGAAAGCTCAGAACAGGTTA 1026
    |||||
DB 113 AAGCAATAGGTTTGTAAACATTTTGTATATACCTTTGAAGGGTTATGATGAACATC 54

QY 1027 AATGCCCTTCT 1037
    |||||
DB 53 TTTCTCTTTAT 43
```

```
Search completed: February 5, 2005, 11:55:46
Job time : 749 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 05:14:35 ; Search time 749 Seconds
(without alignments)
9774.772 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gattcggttcactcgcg.....acaccacaagtaagaattcc 1272

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA:*

1: /cgn2_6/prodata/2/pubpna/US07_PUBCOMB.seq:*

2: /cgn2_6/prodata/2/pubpna/PCT_NEW_PUB.seq:*

3: /cgn2_6/prodata/2/pubpna/US06_NEW_PUB.seq:*

4: /cgn2_6/prodata/2/pubpna/US06_PUBCOMB.seq:*

5: /cgn2_6/prodata/2/pubpna/US07_NEW_PUB.seq:*

6: /cgn2_6/prodata/2/pubpna/PCTUS_PUBCOMB.seq:*

7: /cgn2_6/prodata/2/pubpna/US08_NEW_PUB.seq:*

8: /cgn2_6/prodata/2/pubpna/US08_PUBCOMB.seq:*

9: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*

10: /cgn2_6/prodata/2/pubpna/US09_PUBCOMB.seq:*

11: /cgn2_6/prodata/2/pubpna/US09C_PUBCOMB.seq:*

12: /cgn2_6/prodata/2/pubpna/US09C_NEW_PUB.seq:*

13: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*

14: /cgn2_6/prodata/2/pubpna/US10_PUBCOMB.seq:*

15: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*

16: /cgn2_6/prodata/2/pubpna/US10C_PUBCOMB.seq:*

17: /cgn2_6/prodata/2/pubpna/US10E_PUBCOMB.seq:*

18: /cgn2_6/prodata/2/pubpna/US10F_PUBCOMB.seq:*

19: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*

20: /cgn2_6/prodata/2/pubpna/US10_NEW_PUB.seq:*

21: /cgn2_6/prodata/2/pubpna/US60_NEW_PUB.seq:*

22: /cgn2_6/prodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|---------------------|--------------------|
| 1 | 1272 | 100.0 | 1272 | US-10-077-111-12 | Sequence 12, Appl |
| 2 | 977.2 | 76.8 | 1553 | US-10-077-111-3 | Sequence 3, Appl |
| 3 | 965.8 | 75.9 | 1818 | US-10-077-111-1 | Sequence 1, Appl |
| 4 | 964.2 | 75.8 | 1817 | US-09-971-392-207 | Sequence 207, Appl |
| 5 | 964.2 | 75.8 | 1844 | US-10-037-270-768 | Sequence 768, Appl |
| 6 | 964.2 | 75.8 | 1844 | US-10-117-722-768 | Sequence 768, Appl |
| 7 | 894 | 70.3 | 1773 | US-10-287-218-40 | Sequence 40, Appl |
| 8 | 894 | 70.3 | 1773 | US-10-474-291-40 | Sequence 40, Appl |
| 9 | 890.2 | 70.0 | 1811 | US-10-104-047-1689 | Sequence 1689, Ap |
| 10 | 878.8 | 69.1 | 1996 | US-10-357-930-23179 | Sequence 23179, A |
| 11 | 878.8 | 69.1 | 1996 | US-10-357-930-29046 | Sequence 29046, A |

| | | | | | |
|------|-------|------|------|----------------------|--------------------|
| 12 | 877.8 | 69.0 | 1908 | US-10-077-111-14 | Sequence 14, Appl |
| 13 | 464 | 36.5 | 1901 | US-10-077-111-6 | Sequence 6, Appl |
| 14 | 377.2 | 29.7 | 630 | US-10-077-111-5 | Sequence 5, Appl |
| C 15 | 335 | 26.3 | 446 | US-10-357-930-5029 | Sequence 5029, Ap |
| C 16 | 331.6 | 26.1 | 366 | US-09-864-761-10657 | Sequence 10657, A |
| C 17 | 329.8 | 25.9 | 409 | US-10-357-930-14198 | Sequence 14198, A |
| C 18 | 328.8 | 25.8 | 441 | US-10-357-930-35319 | Sequence 35319, A |
| C 19 | 328.8 | 25.8 | 441 | US-10-357-930-44154 | Sequence 44154, A |
| C 20 | 297 | 23.3 | 297 | US-09-864-761-27306 | Sequence 27306, A |
| C 21 | 179.6 | 14.1 | 466 | US-09-864-761-11052 | Sequence 11052, A |
| C 22 | 176.8 | 13.9 | 180 | US-09-864-761-27687 | Sequence 27687, A |
| C 23 | 154 | 12.1 | 520 | US-10-077-111-8 | Sequence 8, Appl |
| C 24 | 142.6 | 11.2 | 480 | US-09-864-761-13978 | Sequence 13978, A |
| C 25 | 132 | 10.4 | 132 | US-09-864-761-30542 | Sequence 30542, A |
| C 26 | 118.2 | 9.3 | 438 | US-09-918-995-13423 | Sequence 13423, A |
| C 27 | 92 | 7.2 | 428 | US-10-027-632-62257 | Sequence 62257, A |
| C 28 | 92 | 7.2 | 428 | US-10-027-632-62257 | Sequence 62257, A |
| C 29 | 91 | 7.2 | 428 | US-10-027-632-35838 | Sequence 35838, A |
| C 30 | 91 | 7.2 | 428 | US-10-027-632-35838 | Sequence 35838, A |
| C 31 | 87 | 6.8 | 92 | US-10-106-698-3381 | Sequence 3381, Ap |
| C 32 | 76 | 6.0 | 439 | US-10-027-632-185937 | Sequence 185937, A |
| C 33 | 76 | 6.0 | 439 | US-10-027-632-185937 | Sequence 185937, A |
| C 34 | 62.8 | 4.9 | 2511 | US-10-260-238-549 | Sequence 549, Appl |
| C 35 | 62.8 | 4.9 | 2813 | US-10-437-963-64112 | Sequence 64112, A |
| C 36 | 59.6 | 4.7 | 649 | US-10-260-238-4894 | Sequence 4894, Ap |
| C 37 | 59.6 | 4.7 | 1207 | US-10-767-701-12582 | Sequence 12582, A |
| C 38 | 57 | 4.5 | 1839 | US-09-938-842A-1917 | Sequence 1917, Ap |
| C 39 | 57 | 4.5 | 1839 | US-09-938-842A-1917 | Sequence 1917, Ap |
| C 40 | 56.6 | 4.4 | 1265 | US-10-425-115-94246 | Sequence 94246, A |
| C 41 | 56.4 | 4.4 | 604 | US-10-363-345A-4225 | Sequence 4225, Ap |
| C 42 | 56.4 | 4.4 | 604 | US-10-363-345A-4226 | Sequence 4226, Ap |
| C 43 | 56.4 | 4.4 | 604 | US-10-363-345A-5227 | Sequence 5227, Ap |
| C 44 | 56.4 | 4.4 | 604 | US-10-363-345A-5228 | Sequence 5228, Ap |
| C 45 | 56.2 | 4.4 | 1698 | US-10-437-963-27252 | Sequence 27252, A |

ALIGNMENTS

RESULT 1

US-10-077-111-12

Sequence 12, Application US/10077111

Publication No. US20020187492A1

GENERAL INFORMATION:

APPLICANT: Toddertud, C. Gordon

APPLICANT: Finger, Joshua N.

APPLICANT: Rillema, Jill

TITLE OF INVENTION: TBA

FILE REFERENCE: 3053-4114US2

CURRENT APPLICATION NUMBER: US/10/077,111

CURRENT FILING DATE: 2002-02-15

PRIOR APPLICATION NUMBER: 60/294,181

PRIOR FILING DATE: 2001-05-29

PRIOR APPLICATION NUMBER: 60/269,366

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 12

LENGTH: 1272

TYPE: DNA

ORGANISM: HUMAN

FEATURE:

OTHER INFORMATION: RET 16.2 splice variant

US-10-077-111-12

Query Match 100.0%; Score 1272; DB 13; length 1272;

Best local Similarity 100.0%; Pred. No. 0;

Matches 1272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GAATTGGGCTTTACCTGCGGCGACGTCACCGGCGGCGACCTTGAGGCG 60

DB 1 GAATTCGGCTTTACCTGCGGCGGCGACGTCACCGGCGGCGACCTTGAGGCG 60

QY 61 GATCCGCGCGCCCGGCTCTGCAAGCTGTTTTCTTCAAAATPAAGAACATGTAAC 120
DB 61 GATCCGCGCGCCCGGCTCTGCAAGCTGTTTTCTTCAAAATPAAGAACATGTAAC 120
QY 121 TGAATTCACATTAAGTCATGATGTCAGATGTCATCTGCTGCTTCTCTTTTCCC 180
DB 121 TGAATTCACATTAAGTCATGATGTCAGATGTCATCTGCTGCTTCTCTTTTCCC 180
QY 181 TCTTGGCTACTGCTCTGTCGACAAACAAATGCGCTGTATCTGTTACGTACTTAACTG 240
DB 181 TCTTGGCTACTGCTCTGTCGACAAACAAATGCGCTGTATCTGTTACGTACTTAACTG 240
QY 241 AACTGCCACATTCCTCAATGAAGTTTCAATCTATGCTGTCACTGCTGCTTTCTCCC 300
DB 241 AACTGCCACATTCCTCAATGAAGTTTCAATCTATGCTGTCACTGCTGCTTTCTCCC 300
QY 301 CTTCAGAGCATATTTTGGCATGCTGTTCACACATGTTACCACTGTCTTATGAAATCTG 360
DB 301 CTTCAGAGCATATTTTGGCATGCTGTTCACACATGTTACCACTGTCTTATGAAATCTG 360
QY 361 AAAATGACAGATGCTGTCAGATGTCAGACGCTGAGCGCTGTGAGGGTTTGGC 420
DB 361 AAAATGACAGATGCTGTCAGATGTCAGACGCTGAGCGCTGTGAGGGTTTGGC 420
QY 421 AGTTTTCCCGAGCTCCACGCTGTTGGCATCAGGGGCACTGTATGAACTGTGTTTGT 480
DB 421 AGTTTTCCCGAGCTCCACGCTGTTGGCATCAGGGGCACTGTATGAACTGTGTTTGT 480
QY 481 GGAATGCACAGTCATCAAAATTTATATAGATGTGTATGTTAAAGATGGCTCTTGGCG 540
DB 481 GGAATGCACAGTCATCAAAATTTATATAGATGTGTATGTTAAAGATGGCTCTTGGCG 540
QY 541 CATGTGCATTTTCTCCTAATGAGACTTTTGTCACTGCTCTCATGTGTATTTAA 600
DB 541 CATGTGCATTTTCTCCTAATGAGACTTTTGTCACTGCTCTCATGTGTATTTAA 600
QY 601 CAGTGTGGAGTGAATAAATGAGGTGTCTGATAGTGAAGAACATGATCTTGAATTA 660
DB 601 CAGTGTGGAGTGAATAAATGAGGTGTCTGATAGTGAAGAACATGATCTTGAATTA 660
QY 661 CCGTGTGGAGTGTCTCTCAAGCGCTTCTGATGAGAACAAAGCTTCACTTTTTC 720
DB 661 CCGTGTGGAGTGTCTCTCAAGCGCTTCTGATGAGAACAAAGCTTCACTTTTTC 720
QY 721 GACTGGCATGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 780
DB 721 GACTGGCATGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGAT 780
QY 781 TCTTGAAGGCGCACAGAACATCACTGAAGCAATTTACGAAGATTTGTCAGAGAG 840
DB 781 TCTTGAAGGCGCACAGAACATCACTGAAGCAATTTACGAAGATTTGTCAGAGAG 840
QY 841 TCGTCTCAACATGCTTTGTGCAACAGATTTAAAGATCTTGTGTGTAATTTTCAAGAT 900
DB 841 TCGTCTCAACATGCTTTGTGCAACAGATTTAAAGATCTTGTGTGTAATTTTCAAGAT 900
QY 901 ATAAATGATGAGAAAAAGACTGTGTAATCTTACAAAGAAAGTCTGGCTGATGATTTGA 960
DB 901 ATAAATGATGAGAAAAAGACTGTGTAATCTTACAAAGAAAGTCTGGCTGATGATTTGA 960
QY 961 AAAATGATCTTGAAGCTGCTGATGTAAGTGTGAGAAATTTGAAGCTCAGAGCCA 1020
DB 961 AAAATGATCTTGAAGCTGCTGATGTAAGTGTGAGAAATTTGAAGCTCAGAGCCA 1020
QY 1021 AGGTTAATCTCTTTTCTTCAAGAAATCTGATGTAATTTATATGCTCAATACATGAGAAC 1080
DB 1021 AGGTTAATCTCTTTTCTTCAAGAAATCTGATGTAATTTATATGCTCAATACATGAGAAC 1080
QY 1081 TTAATGAAGATCGGTCATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 1140
DB 1081 TTAATGAAGATCGGTCATGTCAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 1140
QY 1141 ATTGATCAGCAAAAGAAAGTACAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 1200

DB 1141 ATTGATCAGCAAAAGAAAGTACAGATGTCAGATGTCAGATGTCAGATGTCAGATGTCAG 1200
QY 1201 TACTTACACCAATATAGACTCTGAATAATGCGCATGATGATGCTGAGACACACAA 1260
DB 1201 TACTTACACCAATATAGACTCTGAATAATGCGCATGATGATGCTGAGACACACAA 1260
QY 1261 AGTAAAGAAATC 1272
DB 1261 AGTAAAGAAATC 1272

RESULT 2
US-10-077-111-3
; Sequence 3, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Toddertud, C. Gordon
; APPLICANT: Flinger, Joshua N.
; APPLICANT: Rillema, J11
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077, 111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294, 181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269, 366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 1553
; TYPE: DNA
; ORGANISM: HUMAN
US-10-077-111-3

Query Match 76.8%; Score 977.2; DB 13; Length 1553;
Best Local Similarity 81.9%; Pred. No. 1.9e-274;
Matches 1265; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

QY 1 GAATTTGGGCTTTCACCTGGCGGCAACGTCACCGGCGGCTGGGCACTTGAAGGG 60
DB 1 GAATTTGGGCTTTCACCTGGCGGCAACGTCACCGGCGGCTGGGCACTTGAAGGG 60
QY 61 GATCCGCGCGCCCGGCTCTGCAAGCTGTTTTCTTCAAAATPAAGAACATGTAAC 120
DB 61 GATCCGCGCGCCCGGCTCTGCAAGCTGTTTTCTTCAAAATPAAGAACATGTAAC 120
QY 121 TGAATTCACATTAAGTCATGATGTCAGATGTCATCTGCTGCTTCTCTTTTCCC 180
DB 121 TGAATTCACATTAAGTCATGATGTCAGATGTCATCTGCTGCTTCTCTTTTCCC 180
QY 181 TCTTGGCTACTGCTCTGTCGACAAACAAATGCGCTGTATCTGTTACGTACTTAACTG 240
DB 181 TCTTGGCTACTGCTCTGTCGACAAACAAATGCGCTGTATCTGTTACGTACTTAACTG 240
QY 241 AACTGCCACATTCCTCAATGAAGTTTCAATCTATGCTGTCACTGCTGCTTTCTCCC 300
DB 241 AACTGCCACATTCCTCAATGAAGTTTCAATCTATGCTGTCACTGCTGCTTTCTCCC 300
QY 301 CTTCAGAGCATATTTTGGCATGCTGTTCACACATGTTACCACTGTCTTATGAAATCTG 360
DB 301 CTTCAGAGCATATTTTGGCATGCTGTTCACACATGTTACCACTGTCTTATGAAATCTG 360
QY 361 AAAATGACAGATGCTGTCAGATGTCAGACGCTGAGCGCTGTGAGGGTTTGGC 420
DB 361 AAAATGACAGATGCTGTCAGATGTCAGACGCTGAGCGCTGTGAGGGTTTGGC 420
QY 421 AGTTTTCCCGAGCTCCACGCTGTTGGCATCAGGGGCACTGTATGAACTGTGTTTGT 480
DB 421 AGTTTTCCCGAGCTCCACGCTGTTGGCATCAGGGGCACTGTATGAACTGTGTTTGT 480
QY 481 GGAATGCACAGTCATCAAAATTTATATAGATGTGTATGTTAAAGATGGCTCTTGGCG 540

Db 481 GGATGACAGCTACAAATTTATATAGATGTGTAAGTGTAAAGTCTCTTGGCGG 540
Qy 541 CATGTGATTTTCTCCATATGGAAGCTTTTTCATCGTGTCTCTCATGTGTATTTAA 600
Db 541 CATGTGATTTTCTCCATATGGAAGCTTTTTCATCGTGTCTCTCATGTGTATTTAA 600
Qy 601 CAGTGTGGATGATTAATGAGTGTCTGATAGTAAAGACATGATCTTGGAAATTA 660
Db 601 CAGTGTGGATGATTAATGAGTGTCTGATAGTAAAGACATGATCTTGGAAATTA 660
Qy 661 CTTGCTGCGATTTTCTTCAAGCAGCTTTCTGATGAGAACTAGTCTTCAAGTTTTC 720
Db 661 CTTGCTGCGATTTTCTTCAAGCAGCTTTCTGATGAGAACTAGTCTTCAAGTTTTC 720
Qy 721 GACTGCGATCATGTGTGATGAGATTTGCCAAGTCAAAATTTGATTTTCTTTACCATTA 780
Db 721 GACTGCGATCATGTGTGATGAGATTTGCCAAGTCAAAATTTGATTTTCTTTACCATTA 780
Qy 781 TCTT----- 784
Db 781 TCTTAGGTTTGAATTAATAATAAGTACATGATGGGCACTGTCTCTGTTCTGG 840
Qy 785 ----- 784
Db 841 CTTGTGCTTTTCCCGTATGAGGAGATGCTAGTCTAGGGTCACTGATGATAGTCTGCA 900
Qy 785 ----- 784
Db 901 TAGTATATGATTAATTAATGATGAGATTAATTAATTAATTAATTAATTAATTAATTA 960
Qy 785 ----- 784
Db 961 TCACAACCTGTGCTTTGACCTTAATACCTTTTACTGCTGTTCAATGACAAA 1020
Qy 785 ----- 784
Db 1021 CAGTGAACATCTGCAATTTGACCTGGAACACTTTGCAAGGAGGACACAAATC 1080
Qy 805 AGCTGAAGCAATTTACGGAAGATGTCAGAGAGTGTCTCAACATGCTTTGTGTGAC 864
Db 1081 AGCTGAAGCAATTTACGGAAGATGTCAGAGAGTGTCTCAACATGCTTTGTGTGAC 1140
Qy 865 AAGATTTAAAGATCTGTGTGATTTTCAAGATGAATTAATGATGAGTGAAGAACTGT 924
Db 1141 AAGATTTAAAGATCTGTGTGATTTTCAAGATGAATTAATGATGAGTGAAGAACTGT 1200
Qy 925 TGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAATGAACTCTTAGAGTGTGCTA 984
Db 1201 TGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAATGAACTCTTAGAGTGTGCTA 1260
Qy 985 GTAAAGTCTGAGGAAATGGAAGCTCAGAGACCAAGTTAAATCCCTTTCTTACAGAA 1044
Db 1261 GTAAAGTCTGAGGAAATGGAAGCTCAGAGACCAAGTTAAATCCCTTTCTTACAGAA 1320
Qy 1045 TTCTGATGATTTTATATGTCATTAATTAAGAACTTAATGAAAGATCCGTCATCGCAT 1104
Db 1321 TTCTGATGATTTTATATGTCATTAATTAAGAACTTAATGAAAGATCCGTCATCGCAT 1380
Qy 1105 CAGATGGCTATTTATGATGAAAGAAAGCAATGGAATTTGATCAGCAAAAGAAAGCTA 1164
Db 1381 CAGATGGCTATTTATGATGAAAGAAAGCAATGGAATTTGATCAGCAAAAGAAAGCTA 1440
Qy 1165 CAAATGCCATGAAATCTTGTCTTCTTCAAGCGGTACTTAACCAAAATGAGACTGTA 1224
Db 1441 CAAATGCCATGAAATCTTGTCTTCTTCAAGCGGTACTTAACCAAAATGAGACTGTA 1500
Qy 1225 AAAATGGCATTAATGATGAGTGTGAGACACCAAAAGTAAAGA 1268
Db 1501 AAAATGGCATTAATGATGAGTGTGAGACACCAAAAGTAAAGA 1544

US-10-077-111-1
; Sequence 1, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Toddard, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jill
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 1053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077, 111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: HUMAN
US-10-077-111-1
Query Match 75.9%; Score 965.8; DB 13; Length 1818;
Best Local Similarity 81.8%; Pred. No. 4.5e-271;
Matches 1253; Conservative 0; Mismatches 2; Indels 276; Gaps 1;
Qy 12 TCACCTGCGGCGACGATGACCCGACCCCGTGGGCACTTGAAGCGGATCCCGGCG 71
Db 49 TCACCTGCGGCGGACGATGACCCGACCCCGTGGGCACTTGAAGCGGATCCCGGCG 108
Qy 72 CCCCCGCTCTGAGGCGTGTCTTCAATTAAGAACTGGAAGTATGATTAACACA 131
Db 109 CCCCCGCTCTGAGGCGTGTCTTCAATTAAGAACTGGAAGTATGATTAACACA 168
Qy 132 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 191
Db 169 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 228
Qy 192 TGCTCTTGGACAAACAAATTCGCTGATGATGATGATGATGATGATGATGATGATGAT 251
Db 229 TGCTCTTGGACAAACAAATTCGCTGATGATGATGATGATGATGATGATGATGATGAT 288
Qy 252 TTCCCATTTGAAGTTTCACTTAATGATGATGATGATGATGATGATGATGATGATGAT 311
Db 289 TTCCCATTTGAAGTTTCACTTAATGATGATGATGATGATGATGATGATGATGATGAT 348
Qy 312 ATTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 371
Db 349 ATTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 408
Qy 372 ATGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 431
Db 409 ATGCTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 468
Qy 432 GACTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 491
Db 469 GACTCCAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 528
Qy 492 TCATCAAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 551
Db 529 TCATCAAAATTAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 588
Qy 552 TCTCTAATGAAAGCTTCTTGTCTCAGTGTCTCTCATGTGTGATTAACAGTGTGGAT 611
Db 589 TCTCTAATGAAAGCTTCTTGTCTCAGTGTCTCTCATGTGTGATTAACAGTGTGGAT 648
Qy 612 GATTAATGAGGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
Db 649 GATTAATGAGGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 708
Qy 672 TTTTCTCAGGCGAGTTTCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 731

```
Db 709 TTTCTTCACAGCCAGTTCTGATGAGAACAAAGTCTTCAGTTTTTTCAGTCGGCATCA 768
Qy 732 TGTGTGAGAGATTGCCAAGTCAAAATTTGGATTGTTCTTTTACCCATATCTT----- 784
Db 769 TGTGTGACAGATTGCCAAGTCAAAATTTGGATTGTTCTTTTACCCATATCTTAGGTTTT 828
Qy 785 ----- 784
Db 829 GAATTAATAATATAAAGTACACTGAGTGGGCACTGTGCTCTGTTGTGCTGTGCTTTT 888
Qy 785 ----- 784
Db 889 TCCCATGATGGGAGATGATCTAGTCTCAGGGTCAGTGTAAATGCTGTCTGATATATGAT 948
Qy 785 ----- 784
Db 949 ACTAATACTGAGATATATCTTACACACTTGCACACTGACACACAGATATGTCACAACTTGT 1008
Qy 785 ----- 784
Db 1009 GCTTTTGACCTAATACCTTTTACTTGTACTGTTCAATGAGCAAAACAGTGAACATC 1068
Qy 785 -----AGCAAGGCGCACAGAACTGAGCTGAAGCA 815
Db 1069 TGGCAATTTGACCTGGAACAATTGGCAAGCAAGGCGCACAGAACTGAGCTGAAGCA 1128
Qy 816 TTACCGAAGATTGGTCAGAGAGAGTGTCTCAACATGGCTTTGTGCAAGATTTTAAA 875
Db 1129 TTTTACCGAAGATTGGTCAGAGAGAGTGTCTCAACATGGCTTTGTGCAAGATTTTAAA 1188
Qy 876 GATCTGTGTGTAATTTTCAAGATGAATTAACATTTGATGAAAAGAACTGTTGAATCTTACA 935
Db 1189 GATCTGTGTGTAATTTTCAAGATGAATTAACATTTGATGAAAAGAACTGTTGAATCTTACA 1248
Qy 936 AAAGAAAGTGTGCTGATGATTTGAAAATTTGAATCTCTTACGACGCTGATGAAAGTCTG 995
Db 1249 AAAGAAAGTGTGCTGATGATTTGAAAATTTGAATCTCTTACGACGCTGATGAAAGTCTG 1308
Qy 996 AGGAAATTTGAAGGCTCAGAGACCAAGTTAAATCCCTTCTTCAGGAAATTCGATGAA 1055
Db 1309 AGGAAATTTGAAGGCTCAGAGACCAAGTTAAATCCCTTCTTCAGGAAATTCGATGAA 1368
Qy 1056 TTTATATGTCCAATTAACCTAAGAACTTATGAAAAGATCCGATCAATCGCATGAGATGGCTAT 1115
Db 1369 TTTATATGTCCAATTAACCTAAGAACTTATGAAAAGATCCGATCAATCGCATGAGATGGCTAT 1428
Qy 1116 TCATATGAAAAAGAGCAATGAAAAATTTGATCAGCAAAAAAGAAAGCTACAGTCCCATG 1175
Db 1429 TCATATGAAAAAGAGCAATGAAAAATTTGATCAGCAAAAAAGAAAGCTACAGTCCCATG 1488
Qy 1176 ACAAAATCTGTGTTCTTCTCAGCGGTACTTACACCAATAGAGACTGTGAAAATGGCCATC 1235
Db 1489 ACAAAATCTGTGTTCTTCTCAGCGGTACTTACACCAATAGAGACTGTGAAAATGGCCATC 1548
Qy 1236 AATAGATGGCTGAGACACACCAAAAGTAAA 1266
Db 1549 AATAGATGGCTGAGACACACCAAAAGTAAA 1579

RESULT 4
US-09-971-392-207
; Sequence 207, Application US/09971392
; Publication No. US20030134283A1
; GENERAL INFORMATION:
; APPLICANT: Peterson, David P.
; APPLICANT: Peterson, Cecelia I.
; APPLICANT: Cooke, Benjamin G.
; TITLE OF INVENTION: GENES REGULATED IN DENDRITIC CELL DIFFERENTIATION
; FILE REFERENCE: PA-0029 US
; CURRENT APPLICATION NUMBER: US/09/971,392
; PRIOR FILING DATE: 2001-10-03
; PRIOR APPLICATION NUMBER: 60/237,652
; PRIOR FILING DATE: 2000-10-03
```

```
; NUMBER OF SEQ ID NOS: 260
; SOFTWARE: PERL Program
; SEQ ID NO 207
; LENGTH: 1817
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Template ID: 158923.9
US-09-971-392-207

Query Match 75.8%; Score 964.2; DB 10; Length 1817;
Best Local Similarity 81.8%; Pred. No. 1.3e-270;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

Qy 12 TCACCTGGCGGAGAGTGAACCGGACCGGCGGACCTTGAAGGGGGAATCCGCGG 71
Db 49 TCACCTGGCGGAGAGTGAACCGGACCGGCGGACCTTGAAGGGGGAATCCGCGG 108
Qy 72 CCCCCGCTCTGAGGCTGTTTTTCTTCAATTAAGAACATGATGAATTCACACA 131
Db 109 CCCCCGCTCTGAGGCTGTTTTTCTTCAATTAAGAACATGATGAATTCACACA 168
Qy 132 TTAGCTGATCATGTGACGATGTCAACTGCTGTGCTTCTCTTTTCCCTTTGGCTACT 191
Db 169 TTAGCTGATCATGTGACGATGTCAACTGCTGTGCTTCTCTTTTCCCTTTGGCTACT 228
Qy 192 TGCTCCTTGGACAAACAAATTTGGCCCTGTACTGTTACGTAAGTTCATGAACTGACAT 251
Db 229 TGCTCCTTGGACAAACAAATTTGGCCCTGTACTGTTACGTAAGTTCATGAACTGACAT 288
Qy 252 TCTCATTTGAAGTTTCACTACTATGCTGTCTCACTGCTGTGTTCTCCCTTCAGACAT 311
Db 289 TCTCATTTGAAGTTTCACTACTATGCTGTCTCACTGCTGTGTTCTCCCTTCAGACAT 348
Qy 312 ATTTTGGCATGCTGTTCACACAGATGATCACTGTCTCTATGGAATCTGAAAATGACAG 371
Db 349 ATTTTGGCATGCTGTTCACACAGATGATCACTGTCTCTATGGAATCTGAAAATGACAG 408
Qy 372 ATGCTGGCAGTATGGAACAGCTAGTGCAAGCCGTAAGAGGTTTGGCAATTTCCCGCA 431
Db 409 ATGCTGGCAGTATGGAACAGCTAGTGCAAGCCGTAAGAGGTTTGGCAATTTCCCGCA 468
Qy 432 GACTCAGCTGTTTGGCATCAGGGGAGCTGATGGAATCTGTGTTTGTGAATGACAG 491
Db 469 GACTCAGCTGTTTGGCATCAGGGGAGCTGATGGAATCTGTGTTTGTGAATGACAG 528
Qy 492 TCATACAAATTAATATGATGTGATGTTTAAAGATGCTCTTGGCGCATGTGCATTT 551
Db 529 TCATACAAATTAATATGATGTGATGTTTAAAGATGCTCTTGGCGCATGTGCATTT 588
Qy 552 TCTCCATATGGAAGCTTCTTGTCACTGGCTCTCAATGATGATTTAAGTGGGAT 611
Db 589 TCTCCATATGGAAGCTTCTTGTCACTGGCTCTCAATGATGATTTAAGTGGGAT 648
Qy 612 GATAAATAGAGTGTCTGATAGTGAAGAAAGACATGATCTTGAATTAACCTGCTCGAT 671
Db 649 GATAAATAGAGTGTCTGATAGTGAAGAAAGACATGATCTTGAATTAACCTGCTCGAT 708
Qy 732 TGTGTGACAGATTGCCAAGTCAAAATTTGGATTGTTCTTTTACCCATATCTTAGGTTTT 784
Db 769 TGTGTGACAGATTGCCAAGTCAAAATTTGGATTGTTCTTTTACCCATATCTTAGGTTTT 828
Qy 785 ----- 784
Db 829 GAATTAATAATATAAAGTACACTGAGTGGGCACTGTGCTCTGTTGTGCTGTGCTTTT 888
Qy 785 ----- 784
```

| | | | |
|----|------|---|------|
| Db | 889 | TCACATGATGGGACAGATCTAGTCTCAGAGGTACAGATTAAGTCTGATAGTATGAT | 948 |
| Qy | 785 | ----- | 784 |
| Db | 949 | ACTAATACTGAGAAATATACTTCAACATTTGACTCAGACACACAGGTATGTCAACATTGT | 1008 |
| Qy | 785 | ----- | 784 |
| Db | 1009 | GCTTTTGACCTTAATACCCTTTTACTTGCTACGTGTTCAATGACAAACAGTGAACATC | 1068 |
| Qy | 785 | -----ACCAAGGCGCAGAAACATCAGCTGAAAGCA | 815 |
| Db | 1069 | TGGCAATTTGACCTGGAACAACCTTTGCCAAGCAGAGACACAGAACATCAGCTGAAGCA | 1128 |
| Qy | 816 | TTTACCGAAGATTGGTCAGAGAGGTCTCTCAACATGCTTTGTGCAACAAGATTTAA | 875 |
| Db | 1129 | TTTACCGAAGATTGGTCAGAGAGATGTCCTCAACATGCTTTGTGCAACAAGATTTAA | 1188 |
| Qy | 876 | GATCTTGTTGGTATTTTTCACATGAATTAACATTTGATGGAAGAACTGTGAATCTTACA | 935 |
| Db | 1189 | GATCTTGTTGGTATTTTTCACATGAATTAACATTTGATGGAAGAACTGTGAATCTTACA | 1248 |
| Qy | 936 | AAAGAAAGTCTGGCTGATGATTTTGAATTAATCTCTAGGACGTGCTAGTAAAGTCTG | 995 |
| Db | 1249 | AAAGAAAGTCTGGCTGATGATTTTGAATTAATCTCTAGGACGTGCTAGTAAAGTCTG | 1308 |
| Qy | 996 | AGGAAATTTGAAGAGCTCAGACCAAGGTTAAATCCCTTTCTTCAGAAATTCCTGATGA | 1055 |
| Db | 1309 | AGGAAATTTGAAGAGCTCAGACCAAGGTTAAATCCCTTTCTTCAGAAATTCCTGATGA | 1368 |
| Qy | 1056 | TTTATATGTCCAAATPACTAGAGAACTTATGAAGATCCGGTTCATGTCATCAGATGGCTAT | 1115 |
| Db | 1369 | TTTATATGTCCAAATPACTAGAGAACTTATGAAGATCCGGTTCATGTCATCAGATGGCTAT | 1428 |
| Qy | 1116 | TCATATGAAAAGGAAGCAATGGAATTTGATCAGCAAAAAGAAACGTACAGTCCCATG | 1175 |
| Db | 1429 | TCATATGAAAAGGAAGCAATGGAATTTGATCAGCAAAAAGAAACGTACAGTCCCATG | 1488 |
| Qy | 1176 | ACAAATCTTGTCTTCTTCAGCGGTACTTACACCAATAGAGACTGTGAAAATGGCCATC | 1235 |
| Db | 1489 | ACAAATCTTGTCTTCTTCAGCGGTACTTACACCAATAGAGACTGTGAAAATGGCCATC | 1548 |
| Qy | 1236 | AATGATGGCTGGAGACACACCCAAAAGTAA | 1266 |
| Db | 1549 | AATGATGGCTGGAGACACACCCAAAAGTAA | 1579 |

RESULT 5
 US-10-037-270-768
 Sequence 768, Application US/10037270
 Publication No. US20030104529A1
 GENERAL INFORMATION:
 APPLICANT: Tang, Y. Tom
 APPLICANT: Liu, Chenghua
 APPLICANT: Asundi, Vinod
 APPLICANT: Zhang, Jie
 APPLICANT: Ren, Feiyan
 APPLICANT: Chen, Rui-hong
 APPLICANT: Zhao, Qing A.
 APPLICANT: Weinman, Tom
 APPLICANT: Xue, Aidong J.
 APPLICANT: Yang, Yonghong
 APPLICANT: Wang, Jian-Rui
 APPLICANT: Zhou, Ping
 APPLICANT: Ma, Yundong
 APPLICANT: Wang, Dunrui
 APPLICANT: Wang, Zhiwei
 APPLICANT: Tillinghaert, John
 APPLICANT: Drmanac, Radoje T.
 TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and
 TITLE OF INVENTION: Polypeptides
 FILE REFERENCE: 784CIP2B
 CURRENT APPLICATION NUMBER: US/10/037.270

[illegible]

Db 845 GAATTAAATATAAAGTACAGTGAAGGCACTGCTCTCTGTTCTGCGTTGCTTTT 904
QY 785 ----- 784
Db 905 TCCCATATGAGGCAAGTCTAGTCTCAGGGCTCAGTGATAAGTCTGATATATGAT 964
QY 785 ----- 784
Db 965 ACTAATACGTGAATATATCTTACACATGTGACGACACAGGTATGTACAACTTGT 1024
QY 785 ----- 784
Db 1025 GCTTTGACCTATATACCTTTTACTGCTAAGTGTCAATGACAAAAGTGAACATC 1084
QY 785 -----AGCAAGGCGCACAGAACATCAGCTGAACAA 815
Db 1085 TGGCAATTTGACCTGGAACACCTTTGCCAAGCAAGGACAGAACATCAGCTGAACAA 1144
QY 816 TTATCCGAAGATTGGTCAGAGAGAGTGTCTCAACATGCTTTGTGCAACAAGATTTAAA 875
Db 1145 TTATCCGAAGATTGGTCAGAGAGAGTGTCTCAACATGCTTTGTGCAACAAGATTTAAA 1204
QY 876 GATCTTTGTGGTATTTTCAAGATGAATTAACATTGATGAGAAAAGAACTGTTGAATCTTACA 935
Db 1205 GATCTTTGTGGTATTTTCAAGATGAATTAACATTGATGAGAAAAGAACTGTTGAATCTTACA 1264
QY 936 AAAGAAAGTGTGCTGATGATTTGAAAATTGAATCTCTAGGACGTCGATGAAGTCTG 995
Db 1265 AAAGAAAGTGTGCTGATGATTTGAAAATTGAATCTCTAGGACGTCGATGAAGTCTG 1324
QY 996 AGGAAAATTGAGAGCTCAGGACCCAGGTTAAATCCCTTCTTCAGGAATTCCTGATGAA 1055
Db 1325 AGGAAAATTGAGAGCTCAGGACCCAGGTTAAATCCCTTCTTCAGGAATTCCTGATGAA 1384
QY 1056 TTTATATGTCATTAATCTAGAGAACTTATGAAAGATCCGGTCATGCACTAGATGCTAT 1115
Db 1385 TTTATATGTCATTAATCTAGAGAACTTATGAAAGATCCGGTCATGCACTAGATGCTAT 1444
QY 1116 TCATATATAAAGAGAAAGCAATGAGAAATTTGATCAGAAAAGAAAGCAATGAGTCCCATG 1175
Db 1445 TCATATATAAAGAGAAAGCAATGAGAAATTTGATCAGAAAAGAAAGCAATGAGTCCCATG 1504
QY 1176 ACAAAATCTTGTCTTCTTCAGCGGTACTTACACCAATATGAGACTTGAAATATGCCATC 1235
Db 1505 ACAAAATCTTGTCTTCTTCAGCGGTACTTACACCAATATGAGACTTGAAATATGCCATC 1564
QY 1236 AATATGATGCTGAGAGACACACCAAAAGTAAA 1266
Db 1565 AATATGATGCTGAGAGACACACCAAAAGTAAA 1595

RESULT 6
US-10-117-722-768
; Sequence 768, Application US/10117722
; Publication No. US20030219744A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Adundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2BCIP
; CURRENT APPLICATION NUMBER: US/10/117, 722
; PRIOR FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/620,312
; PRIOR FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21

; NUMBER OF SEQ ID NOS: 1104
; SOFTWARE: pc_fl_genes Version 1.0
; SEQ ID NO 768
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164) .. (1594)
US-10-117-722-768

Query Match 75.8%; Score 964.2; DB 17; Length 1844;
Best Local Similarity 81.8%; Pired. No. 1.3e-270;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

QY 12 TCACCTGCGGCGGACGAGACCCGACCGCCCGGAGCATTGAAGCGGATCCCGCG 71
Db 65 TCACCTGCGGCGGACGAGACCCGACCGCCCGGAGCATTGAAGCGGATCCCGCG 124
QY 72 CCCCCGCTCTGACAGGCTGTTTCTTCAATTAAGAACATGGTGAACGTATTCACACA 131
Db 125 CCCCCGCTCTGACAGGCTGTTTCTTCAATTAAGAACATGGTGAACGTATTCACACA 184
QY 132 TTAGCTGATATGTTGACGATGTCAACTGCTGTGCTTCTCTTTCCCTTGGCTACT 191
Db 185 TTAGCTGATATGTTGACGATGTCAACTGCTGTGCTTCTCTTTCCCTTGGCTACT 244
QY 192 TGTCTCTGGAACAAACAAATGCGCTGTACTGTTACGTGAACCTTACGAACTGCAACAT 251
Db 245 TGTCTCTGGAACAAACAAATGCGCTGTACTGTTACGTGAACCTTACGAACTGCAACAT 304
QY 252 TCTCCATTGAAGTTTCATACCTATGCTGTCCACTGCTGTGTTCTCCCTTCAGACAT 311
Db 305 TCTCCATTGAAGTTTCATACCTATGCTGTCCACTGCTGTGTTCTCCCTTCAGACAT 364
QY 312 ATTTTGGCATGTTTCAACAGATGTGACCACTGTCTATGAAATCTGAAAATGACAG 371
Db 365 ATTTTGGCATGTTTCAACAGATGTGACCACTGTCTATGAAATCTGAAAATGACAG 424
QY 372 ATGCTGGAGTGAATGGAACAGCTAGTGGCAGCCGTGAGGGTTGGCCATTTCCCA 431
Db 425 ATGCTGGAGTGAATGGAACAGCTAGTGGCAGCCGTGAGGGTTGGCCATTTCCCA 484
QY 432 GACTCCACGTTTGGCATCAGGGGACGTGATGAACTGTGTTTGTGAAATGACAG 491
Db 485 GACTCCACGTTTGGCATCAGGGGACGTGATGAACTGTGTTTGTGAAATGACAG 544
QY 492 TCATCAAAATTAATATGATGTGTAGTATTAAGATGGCTCCTTGGCGGATGTGATTT 551
Db 545 TCATCAAAATTAATATGATGTGTAGTATTAAGATGGCTCCTTGGCGGATGTGATTT 604
QY 552 TCTCCTAATGGAAGCTTCTTGTGACATGCTCCTCATATGTGTGATTTAACAGTGGGAT 611
Db 605 TCTCCTAATGGAAGCTTCTTGTGACATGCTCCTCATATGTGTGATTTAACAGTGGGAT 664
QY 612 GATAAATAGAGTGTGCTAGTGAAGAAAGCAGATGTTTGAATTAACCTGCTCGAT 671
Db 665 GATAAATAGAGTGTGCTAGTGAAGAAAGCAGATGTTTGAATTAACCTGCTCGAT 724
QY 732 TGTGTGACAGATGCGCAAGTCAAAATTTGATGTGTTTCTTTTACCATATCTT----- 784
Db 785 TGTGTGACAGATGCGCAAGTCAAAATTTGATGTGTTTCTTTTACCATATCTTAAAGTTT 844
QY 785 ----- 784
Db 845 GAATTAAATATAAAGTACAGTGAAGGCACTGCTCTCTGTTCTGCGTTGCTTTT 904
QY 785 ----- 784

| QY | DB | Sequence | Position |
|----|------|---|----------|
| QY | 905 | UCCCATGATGGCAGATGCTAGTCTCAGGGCTAGTGGATTAATCTGTCTATATATATGAT | 964 |
| QY | 785 | ----- | 784 |
| QY | 965 | ACTAATACTGAGAAATATATCTTACACATTTGACTCAGACACACAGGATATGTCAAACTTGT | 1024 |
| QY | 785 | ----- | 784 |
| QY | 1025 | GCTTTGCACTTAATACCTTTTACTTGCTTACTGTTCAATGACAAAACAGTGAACATC | 1084 |
| QY | 785 | -----AGCAAGCCGACAGAAATCATGAGTGAAGCA | 815 |
| QY | 1085 | TGCGAATTTGACCTGGAAACAATTGCCAAGCAGAGACAGACAGACATAGCTGAAGCA | 1144 |
| QY | 816 | TTTACCGAAGATGGTCTCAGAGAGGTGCTCAACATGGCTTTGTGCAACAAGATTTAAA | 875 |
| QY | 1145 | TTTACCGAAGATGGTCTCAGAGAGGTGCTCAACATGGCTTTGTGCAACAAGATTTAAA | 1204 |
| QY | 876 | GATCTTTGTTGTTATTTTCAAGATGAATTAACATTTGATGAAAAAAGTCTGTTGAATCTTACA | 935 |
| QY | 1205 | GATCTTTGTTGTTATTTTCAAGATGAATTAACATTTGATGAAAAAAGTCTGTTGAATCTTACA | 1264 |
| QY | 936 | AAAGAAAGTCTGGCTGATGATTTGAAAATTTGATCTCTAGACCTGCTGTAAGTCTG | 995 |
| QY | 1285 | AAAGAAAGTCTGGCTGATGATTTGAAAATTTGATCTCTAGACCTGCTGTAAGTCTG | 1324 |
| QY | 996 | AGAAAATTTGAAGAGCTCAGACCAAGGTTAAATCCCTTTCTTCAGAAATTCCTGATGAA | 1055 |
| QY | 1325 | AGAAAATTTGAAGAGCTCAGACCAAGGTTAAATCCCTTTCTTCAGAAATTCCTGATGAA | 1384 |
| QY | 1056 | TTTATATGTCCTCAATTAATCTAGAACTTATGAAAAGTCCGCTCATGCGATGAGTGGCTAT | 1115 |
| QY | 1385 | TTTATATGTCCTCAATTAATCTAGAACTTATGAAAAGTCCGCTCATGCGATGAGTGGCTAT | 1444 |
| QY | 1116 | TCATATGAAAAGAGCAATGGAATTTGGATCAGCAAAAAGAAACGTCAAGTCCCATG | 1175 |
| QY | 1445 | TCATATGAAAAGAGCAATGGAATTTGGATCAGCAAAAAGAAACGTCAAGTCCCATG | 1504 |
| QY | 1176 | ACAAATCTGTTCTTCTTCAGCGGTACTTACACCAATATGAGCTCTGAAAATGGCCATC | 1235 |
| QY | 1505 | ACAAATCTGTTCTTCTTCAGCGGTACTTACACCAATATGAGCTCTGAAAATGGCCATC | 1564 |
| QY | 1236 | AATAGATGCTGAGACACACCAAAAGTTAA | 1266 |
| QY | 1565 | AATAGATGCTGAGACACACCAAAAGTTAA | 1595 |
| DB | | | |

[illegible]

Db 320 GATGCTGGCAGTATGGAACAGCCTTAGTGGCAGCCCTGTGAAGGTTTGCACGTTTTCCCTCC 379
 Qy 431 AGACTCCACGCTGTTTTGGCATACAGGGCAGCCTGATGGAACCTGCTTTTGGAAATGCACA 490
 Db 380 AGACTCCACGCTGTTTGGCATACAGGGCAGCCTGATGGAACCTGCTTTTGGAAATGCACA 433
 Qy 491 GTCATACAAATTATATAGATGTTGATGTAGTGTAAAGATGGCTCTTGGCGGCGATGCAATT 550
 Db 440 GTCATACAAATTATATAGATGTTGATGTAGTGTAAAGATGGCTCTTGGCGGCGATGCAATT 499
 Qy 551 TTCTCCTAATGAGAAGCTTCTTTGTCACTGGCTTCTCATGTGTGATTTTAAACAGTGTGGGA 610
 Db 500 TTCTCCTAATGAGAAGCTTCTTTGTCACTGGCTTCTCATGTGTGATTTTAAACAGTGTGGGA 555
 Qy 611 TGATAAATAGAGGTGTCTGCATATGTGAATAAGACATGATCTTGGAAATTACCTGTGGGA 670
 Db 560 TGATAAATAGAGGTGTCTGCATATGTGAATAAGACATGATCTTGGAAATTACCTGTGGGA 615
 Qy 671 TTTTTCCTTCACAGCCAGTTTCTGATGTGAATAAGAGTCTTCAGTTTTTTCGACTGGCATC 733
 Db 620 TTTTTCCTTCACAGCCAGTTTCTGATGTGAATAAGAGTCTTCAGTTTTTTCGACTGGCATC 679
 Qy 731 ATGTGTCAGAGATTGCAGATCAAAATTGGAATGTGTTCTTTTACCCTATATCTT----- 788
 Db 680 ATGTGTCAGAGATTGCAGATCAAAATTGGAATGTGTTCTTTTACCCTATATCTTAAAGTTT 733

[illegible]

| | | | |
|----|------|--|------|
| Oy | 785 | ----- | 784 |
| Db | 800 | TTCCCATGATGGGCAGATGCTAGTCTCAGGGTCAGTGGATTAAGTCTGCATAGTATATGA | 859 |
| Oy | 785 | ----- | 784 |
| Db | 860 | TACTAATACAGAAATATACTTTCACATGTGACTGCAGCACACAGGTATGCACAACTTG | 919 |
| Oy | 785 | ----- | 784 |
| Db | 920 | TGCTTTTGACCTTAATACCCTTTTACTTGCTACTGCTTCAAATGACAAACAGTGAAAT | 979 |
| Oy | 785 | -----AGCAAGGCGCACAGAACATCAGCTGAAGCA | 814 |
| Db | 980 | CTGGCAATTTGACCTGAAACACTTTGGCCAAAGCAAGGAGCACAAGAACATCAGCTGAAGCA | 1039 |
| Oy | 815 | ATTTAATCCGAATTTGTCAGAGAGAGTGTCTCAACATGCGTTTGTGCACAAATTTTAA | 874 |
| Db | 1040 | ATTTAATCCGAATTTGTCAGAGAGAGTGTCTCAACATGCGTTTGTGCACAAATTTTAA | 1099 |
| Oy | 875 | AGATCTTGTTGGTATTTTCAAGATGAATACATGATGGAAGAAAGACGTGTGAATCTTAC | 934 |
| Db | 1100 | AGATCTTGTTGGTATTTTCAAGATGAATACATGATGGAAGAAAGACGTGTGAATCTTAC | 1159 |
| Oy | 935 | AAAAGAAAGTCTGCGTGAATTTGAAAATTGAATCTTACAGATCGCGTAAAGTCT | 994 |
| Db | 1160 | AAAAGAAAGTCTGCGTGAATTTGAAAATTGAATCTTACAGATCGCGTAAAGTCT | 1219 |
| Oy | 995 | GAGGAAATTTGAAGAGCTCAGGACCCAAAGTTAAATCCCTTCTTCAGGAAATTCCTGATGA | 1054 |
| Db | 1220 | GAGGAAATTTGAAGAGCTCAGGACCCAAAGTTAAATCCCTTCTTCAGGAAATTCCTGATGA | 1279 |
| Oy | 1055 | ATTATATATGCTCAATATCTAGAACTTATGAAAGATCCGGTCATGCGATCAGATGGCTA | 1114 |
| Db | 1280 | ATTATATATGCTCAATATCTAGAACTTATGAAAGATCCGGTCATGCGATCAGATGGCTA | 1339 |
| Oy | 1115 | TTCAATATGAAGAAGCAATGGAATAATGGATCAGAAAAGAAAGGTACAGATGCCAT | 1174 |
| Db | 1340 | TTCAATATGAAGAAGCAATGGAATAATGGATCAGAAAAGAAAGGTACAGATGCCAT | 1399 |
| Oy | 1175 | GACAAATCTTGTTCTTCCCTTACGGGTACTTACACCAATAGGAATCTGAAAATGGCCAT | 1234 |

| | | | |
|----|-----|---|-----|
| OY | 86 | GGCGTTTTCCTCAAAATTAAGAACATGGTAAATCTAATTCACATAGTATGTCATGG | 145 |
| Db | 125 | GGCTGTTTTCTTCAAAATTAAGAACATGGTAAATCTAATTCACATAGTATGTCATGG | 184 |
| OY | 146 | TGACGATGTCAACTGCTGTGCTTCTCTTTTCCCTCTTGAGCTACCTTGCTCTTGACAA | 205 |
| Db | 185 | TGAGGATGTCAACTGCTGTGCTTCTCTTTTCCCTCTTGAGCTACCTTGCTCTTGACAA | 244 |
| OY | 206 | AACAAATTCGCTGACTGTGCTTGAAGTAACTTACTGAACTGGCAATTCGATGGAAGTT | 265 |
| Db | 245 | AACAAATTCGCTGACTGTGCTTGAAGTAACTTACTGAACTGGCAATTCGATGGAAGTT | 304 |
| OY | 266 | TCATACCTATGCTGTGCTGCACTGCTGTCTTCTCCCTTCAGAGACATATTTTGGATCGTG | 325 |
| Db | 305 | TCATACCTATGCTGTGCTGCACTGCTGTCTTCTCCCTTCAGAGACATATTTTGGATCGTG | 364 |
| OY | 326 | TTCAACAGATGATACCACTGTCTCTTAATGGAATATCTGAAATATGGAACAGATGCGGGAGAT | 385 |
| Db | 365 | TTCAACAGATGATACCACTGTCTCTTAATGGAATATCTGAAATATGGAACAGATGCGGGAGAT | 424 |
| OY | 386 | GGAACACGCTTAGTGGGACGCTGTGAGGGGTTTGGCAGTTTCCCGACACTCCACGTGTTT | 445 |
| Db | 425 | GGAACACGCTTAGTGGGACGCTGTGAGGGGTTTGGCAGTTTCCCGACACTCCACGTGTTT | 484 |
| OY | 446 | GGCATTCAGGGGACACTGATGGAATCTGTGGTTTTGTGGATATGCAACATATCAAAATTATA | 505 |
| Db | 485 | GGCATTCAGGGGACACTGATGGAATCTGTGGTTTTGTGGATATGCAACATATCAAAATTATA | 544 |
| OY | 506 | TAGATGTGATGTATTAAAGATGGCTCTTGAGCGGCAATGTGCATTTTCTCCTTAATGGAAG | 565 |
| Db | 545 | TAGATGTGATGTATTAAAGATGGCTCTTGAGCGGCAATGTGCATTTTCTCCTTAATGGAAG | 604 |
| OY | 566 | CTTCTTTGTCACTGGCTCTCTCATATGTGATTTAAACAGTGTGGATGTATAAATGAGGTG | 625 |
| Db | 605 | CTTCTTTGTCACTGGCTCTCTCATATGTGATTTAAACAGTGTGGATGTATAAATGAGGTG | 664 |
| OY | 626 | TCTGCATGTATAAAGACATATGATCTTGGAAATTAACCGTGTGCATTTTCTTCACAGCC | 685 |
| Db | 665 | TCTGCATGTATAAAGACATATGATCTTGGAAATTAACCGTGTGCATTTTCTTCACAGCC | 724 |
| OY | 686 | AGTTTCTGATGAGAACAAAGGCTTTCAGTTTTCGACTGGCATCATGTGCTCAGGATTG | 745 |

Db 725 AGTTTCGTAGTGAGAACAGAGCTCTTCAGTTTTCGACGTGCATCATGTGTCAGGATTG 784
QY 746 CCAAGTCAAAATTTGGATTGTTCTTTTACCCATATCTT----- 784
Db 785 CCAAGTCAAAATTTGGATTGTTCTTTTACCCATATCTTAAAGTTTAAATATA 844
QY 785 ----- 784
Db 845 AAGTACACTGATGAGGCACTGCTCCTGTTTGGCTTGTGCTTTTCCCATGATGGGA 904
QY 785 ----- 784
Db 905 GATGCTAGTCTCAGGGCTCAGTGATTAAGTCTGTCAATAGTATGATATCTAATCTGAGAA 964
QY 785 ----- 784
Db 965 TATACTTCACACATTTGACTGACACACAGATATGTCAACAATTGCTTTTGCACTAA 1024
QY 785 ----- 784
Db 1025 TACCTTTTACTGCTACTGTTCAATGACAAAACAGTGAACATCTGCAATTTGACT 1084
QY 785 -----AGCAAGCGCCACAGAAACATCAAGCTGAAGAAATTTACGAAAGTTG 829
Db 1085 GGAACAACCTTGGCCAGAGAGACAGACAGAACTCAGCTGAAGCAATTTTCCGAAAGATTG 1144
QY 830 GTGAGAGAGAGTGTCTCAACATGAGCTTTGTGCAAGATTTTAAAGATCTTGTGGTAT 889
Db 1145 GTGAGAGAGAGATGTCTCAACATGAGCTTTGTGCAAGATTTTAAAGATCTTGTGGTAT 1204
QY 890 TTTCAAGATGATTAACATTTGATGAGAAAAGAACTGTTGAATCTTACAAAAGAAAGTCTGC 949
Db 1205 TTTCAAGATGATTAACATTTGATGAGAAAAGAACTGTTGAATCTTACAAAAGAAAGTCTGC 1264
QY 950 TGTGATTTTAAATTTGAATCTCTAGAGACGTCGTAAGTAAAGTGTGAGAAAATTTGAAGA 1009
Db 1265 TGTGATTTTAAATTTGAATCTCTAGAGACGTCGTAAGTAAAGTGTGAGAAAATTTGAAGA 1324
QY 1010 GCTCAGAGACCAAGTTAAATCCCTTTCTTCAGGAATCTCTGATGATTTTATGTCGAAT 1069
Db 1325 GCTCAGAGACCAAGTTAAATCCCTTTCTTCAGGAATCTCTGATGATTTTATGTCGAAT 1384
QY 1070 AACTAGAGAACTTATGAGAAATCCGTCATTCGATCAGATGAGCTATTCATATGAGAAAAGA 1129
Db 1385 AACTAGAGAACTTATGAGAAATCCGTCATTCGATCAGATGAGCTATTCATATGAGAAAAGA 1444
QY 1130 AGCAATGAGAAATTTGATTCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCT 1189
Db 1445 AGCAATGAGAAATTTGATTCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCT 1504
QY 1190 TCCTTCAGCGTACTTACACAAATAGAGACTGTGAAATAGGCCATCAATAGATGGCTGA 1249
Db 1505 TCCTTCAGCGTACTTACACAAATAGAGACTGTGAAATAGGCCATCAATAGATGGCTGA 1564
QY 1250 GACACACCAAAAGTAAA 1266
Db 1565 GACACACCAAAAGTAAA 1581

RESULT 10
US-10-357-930-23179
; Sequence 23179, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04

; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 23179
; LENGTH: 1996
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2,3,1987,1988,1989,1990,1991,1992,1993,1994,
; LOCATION: 1995,1996
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-23179
Query Match 69.1%; Score 878.8; DB 18; Length 1996;
Best Local Similarity 80.7%; Pred. No. 1.3e-245;
Matches 1166; Conservative 0; Mismatches 2; Indels 276; Gaps 1;
QY 99 CAATTAAGAACATGTGTAAGTCAACATGATTCACACATTAAGTCAATGATCATGTCAGATGTCAC 158
Db 4 CAATTAAGAACATGTGTAAGTCAACATGATTCACACATTAAGTCAATGATCATGTCAGATGTCAC 63
QY 159 TGTGTGCTTCTCTCTTTTCCCTCTTGGCTACTTGTCTTGGACAAAACATTCGCTG 218
Db 64 TGTGTGCTTCTCTCTTTTCCCTCTTGGCTACTTGTCTTGGACAAAACATTCGCTG 123
QY 219 TACTGTTACGTACTTACTGTAAGTCAACATTCGTAATGATTCATGATTCATGATTCATGATTC 278
Db 124 TACTGTTACGTACTTACTGTAAGTCAACATTCGTAATGATTCATGATTCATGATTCATGATTC 183
QY 279 GTCCACTGCTGCTGTTTCTCCCTTCAGAGACATATTTGGCATTCGTTCAACAGATGCT 338
Db 184 GTCCACTGCTGCTGTTTCTCCCTTCAGAGACATATTTGGCATTCGTTCAACAGATGCT 243
QY 339 ACCACTGCTCTATGAAATCTGAAAATGACAGATGCTGCACTGAGTGAACAGCTTACT 398
Db 244 ACCACTGCTCTATGAAATCTGAAAATGACAGATGCTGCACTGAGTGAACAGCTTACT 303
QY 399 GGCAGCCCTGTGAGGGTTTGGCAGTTTCCCGAGACTCCACGATGTTGGCATCAGGGGA 458
Db 304 GGCAGCCCTGTGAGGGTTTGGCAGTTTCCCGAGACTCCACGATGTTGGCATCAGGGGA 363
QY 459 GCTGATGAACTGTGTTTGTGGAATGACAGTCAATAAAATATATAGATGTGTAGT 518
Db 364 GCTGATGAACTGTGTTTGTGGAATGACAGTCAATAAAATATATAGATGTGTAGT 423
QY 519 GTTAAAGATGCTCTTGGCGGAGTGCATTTTCTCTAATGAAAGCTTTTGTCTACT 578
Db 424 GTTAAAGATGCTCTTGGCGGAGTGCATTTTCTCTAATGAAAGCTTTTGTCTACT 483
QY 579 GGTCTCTCATGTGTAATTTAACAAGTGTGGAATGAAAATGAGTGTCTGCATATGAA 638
Db 484 GGTCTCTCATGTGTAATTTAACAAGTGTGGAATGAAAATGAGTGTCTGCATATGAA 543
QY 639 AAGGACATGATTTGAAATTAACCTGCTGATTTTCTTCACAGCAAGTTTCTGATGA 698
Db 544 AAGGACATGATTTGAAATTAACCTGCTGATTTTCTTCACAGCAAGTTTCTGATGA 603
QY 699 GAAACAAGTCTTCAAGTTTTCGACATGCAATGCTCAGATTTGCAAGTCAAAAT 758

Db 604 GAAAGAGCTTCACTTTTTCAGCTGCGATCATGTGCTGAGATTGCCAAGTCAAAATT 663
Qy 759 TGAATGTTCTTTTACCATATCTT----- 784
Db 664 TGAATGTTCTTTTACCATATCTTGTAGTTTGAATTAATAATAAAGTACAGT 723
Qy 785 ----- 784
Db 724 GGGAGCTGTGCTCTGTTCTGGCTTGTGCTTTTCCATGATGGGAGAGTGTAGTCTCA 783
Qy 785 ----- 784
Db 784 GGGTCACTGATAGTCTGTATAGTATAGTACTAATAGTGAATATATCTTACACA 843
Qy 785 ----- 784
Db 844 TTGACTGACGACACAGGATATGTACAACATTGTGCTTTTGACCTAATACCTTTTACTT 903
Qy 785 ----- 784
Db 904 GCTACTGTTCAATGACAAAGAGTGAACATCTGGCAATTTGACCTGGAAACACTTTC 963
Qy 785 --AGCAAGGCGCAGAAACATCAGCTGAGCAATTTACCGAAGTTGCTCAGAGAGTCT 842
Db 964 CAACCAAGGCGCAGAAACATCAGCTGAGCAATTTACCGAAGTTGCTCAGAGAGAT 1023
Qy 843 GTCTCAATGCTGTTGTGCAAGATTTAAAGTCTGTGATTTTCAAGATGAT 902
Db 1024 GTCTCAATGCTGTTGTGCAAGATTTAAAGTCTGTGATTTTCAAGATGAT 1083
Qy 903 AACATGATGAAAGAACTGTTGATCTTCAAAAGAAAGTCTGCTGATGATTTGAA 962
Db 1084 AACATGATGAAAGAACTGTTGATCTTCAAAAGAAAGTCTGCTGATGATTTGAA 1143
Qy 963 ATTGAATCTTGAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1022
Db 1144 ATTGAATCTTGAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1203
Qy 1023 GTTAAATCCCTTCTTCAAGAAATCCGATGATGATGATGATGATGATGATGATGATGAT 1082
Db 1204 GTTAAATCCCTTCTTCAAGAAATCCGATGATGATGATGATGATGATGATGATGATGAT 1263
Qy 1083 ATGAAGATCCGCTCATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1142
Db 1264 ATGAAGATCCGCTCATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1233
Qy 1143 TGAATCAGCAAAAGAAAGTGAAGTCCATGACAAATCTTCTTCTTCAAGCGTA 1202
Db 1224 TGAATCAGCAAAAGAAAGTGAAGTCCATGACAAATCTTCTTCTTCAAGCGTA 1383
Qy 1203 CTTTACCAATAGGATCTTGAATTAATGGCCATCAATAGATGGCTGGAGACACCAAAAG 1262
Db 1384 CTTTACCAATAGGATCTTGAATTAATGGCCATCAATAGATGGCTGGAGACACCAAAAG 1443
Qy 1263 TAAA 1266
Db 1444 TAAA 1447

RESULT 11
US-10-357-930-29046
; Sequence 29046, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04

PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 29046
; LENGTH: 1996
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: 1, 2, 3, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994,
; LOCATION: 1995, 1996
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-29046
Query Match 69.1%; Score 878.8; DB 18; Length 1996;
Best Local Similarity 80.7%; Pred. No. 1.3e-245; Indels 276; Gaps 1;
Matches 1166; Conservative 0; Mismatches 2;
Qy 99 CAATTAAGAAAGATGTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 158
Db 4 CAATTAAGAAAGATGTAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 63
Qy 159 TGCTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 218
Db 64 TGCTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 123
Qy 219 TACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 278
Db 124 TACTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 183
Qy 279 GTCCAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 338
Db 184 GTCCAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 243
Qy 339 ACCAATGCTTATGAAATGATGAAATGATGAAATGATGAAATGATGAAATGATGAAATGAT 398
Db 244 ACCAATGCTTATGAAATGATGAAATGATGAAATGATGAAATGATGAAATGATGAAATGAT 303
Qy 399 GGCAGCCCTGTGAGGTTTGGCACTTTTCCCGAGCTTCCAGCTTCCAGCTTCCAGCTTCC 458
Db 304 GGCAGCCCTGTGAGGTTTGGCACTTTTCCCGAGCTTCCAGCTTCCAGCTTCCAGCTTCC 363
Qy 459 GCTGATGAACTGTGTTTGTGATGAAATGCAAGTCAATCAATTAATTAATTAATTAATTA 518
Db 364 GCTGATGAACTGTGTTTGTGATGAAATGCAAGTCAATCAATTAATTAATTAATTAATTA 423
Qy 519 GTTAAAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 578
Db 424 GTTAAAGTGTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 483
Qy 579 GGGTCTCTATGTGATGATTAACAGTGTGGATGATTAATTAATTAATTAATTAATTAATTA 638
Db 484 GGGTCTCTATGTGATGATTAACAGTGTGGATGATTAATTAATTAATTAATTAATTAATTA 543
Qy 639 AAAGCAATGATCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 698
Db 544 AAAGCAATGATCTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 603
Qy 699 GAAAGATCTTCAATTTTTCAGCTGCGATCATGTGATGATGATGATGATGATGATGATGAT 758

Db 604 GAACAAAGTCTTCAAGTTTTCAGCTGGCATCATGTGCTGAGATTGCCAAGTCAAAATT 663
QY 759 TGGATTGTTTCTTTTACCATATCTT----- 784
Db 664 TGGATTGTTTCTTTTACCATATCTTGTGTTGAATTAATAATAGTACAGT 723
QY 785 ----- 784
Db 724 GGGCAGTGTGCTCTGTCTTGGCTTGTCTTCCATGATGGGAGATGTAGTCTCA 783
QY 785 ----- 784
Db 784 GGGCAGTGTGATAGTCTGATAGTATGATAGTAACTAGAAATATATCTTACACA 843
QY 785 ----- 784
Db 844 TTGACTCAGACACACAGATATGTACAACTTGTGCTTTTGACCTATACCTTTTACT 903
QY 785 ----- 784
Db 904 GCTACTGTCTCAATGACAAACAGTGAACATCTGCAATTTGACCTGAAACACTTTC 963
QY 785 --AGCAAGGGGCAAGAACTCATGCTGAGCAATTTTACCGAAGTTGGTGAAGAGTCTC 842
Db 964 CAAGCAAGGGGCAAGAACTCATGCTGAGCAATTTTACCGAAGTTGGTGAAGAGTCTC 1023
QY 843 GTCTCAACATGAGCTTTGTGACAAAGATTTTAAAGATCTTGTGTATTTTCAAGATGAT 902
Db 1024 GTCTCAACATGAGCTTTGTGACAAAGATTTTAAAGATCTTGTGTATTTTCAAGATGAT 1083
QY 903 AACATGATGAAAGAAAGAACTGTTGATCTTACAAAGAAAGTCTGCTGATGATTTGAA 962
Db 1084 AACATGATGAAAGAAAGAACTGTTGATCTTACAAAGAAAGTCTGCTGATGATTTGAA 1143
QY 963 ATTGAATCTTGAAGCTGCGCTAGTAAAGTCTGAGAAATTTGAAGCTCAGAACCAAG 1022
Db 1144 ATTGAATCTTGAAGCTGCGCTAGTAAAGTCTGAGAAATTTGAAGCTCAGAACCAAG 1203
QY 1023 GTTAAATCCCTTCTTTCAGGAATTCCTGATGAATTTTATATGTCATATGATGAGAACTT 1082
Db 1204 GTTAAATCCCTTCTTTCAGGAATTCCTGATGAATTTTATATGTCATATGATGAGAACTT 1263
QY 1083 ATGAAAGATCCGCTCATGCGATCAGATGGCTATTCATATGAAAGAAAGCAATGGAAT 1142
Db 1264 ATGAAAGATCCGCTCATGCGATCAGATGGCTATTCATATGAAAGAAAGCAATGGAAT 1223
QY 1143 TGGATCAGCAAAAGAAAGCAATGTCATGCAATCTTGTCTTCTTCAAGCGCTA 1202
Db 1324 TGGATCAGCAAAAGAAAGCAATGTCATGCAATCTTGTCTTCTTCAAGCGCTA 1383
QY 1203 CTACACCAAAATGAGCTCTGAAATATGGCATCAATGATGGCTGAGACACCAAAAG 1262
Db 1384 CTACACCAAAATGAGCTCTGAAATATGGCATCAATGATGGCTGAGACACCAAAAG 1443
QY 1263 TAAA 1266
Db 1444 TAAA 1447

RESULT 12
US-10-077-111-14
; Sequence 14, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, J111
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/269,181
; PRIOR FILING DATE: 2001-05-29

QY 69.0%; Score 877.8; DB 13; Length 1908;
Best Local Similarity 77.9%; Pred. No. 2.5e-245;
Matches 1253; Conservative 0; Mismatches 2; Indels 354; Gaps 2;

QY 12 TCACCTGGCGGAGACGTGACCCGACCCGCGGACCTTGAAGGGGATCCCGCGG 71
Db 37 TCACCTGGCGGAGACGTGACCCGACCCGCGGACCTTGAAGGGGATCCCGCGG 96
QY 72 CCCCCGCTCTGAGGCTGTTTCTTCAAAATPAAAGATGGTGAAGTATTCACACA 131
Db 97 CCCCCGCTCTGAGGCTGTTTCTTCAAAATPAAAGATGGTGAAGTATTCACACA 156
QY 132 TTAGCTGATCAGTGTGACAGATGCAACCTGCTGCTTCTTCTTCTTCTTCTTCT 191
Db 157 TTAGCTGATCAGTGTGACAGATGCAACCTGCTGCTTCTTCTTCTTCTTCTTCT 216
QY 192 TGGTCTTGGACAAACAAATGCGCTGTACTCTGTTACGTAACCTTACCTGAACCTG 251
Db 217 TGGTCTTGGACAAACAAATGCGCTGTACTCTGTTACGTAACCTTACCTGAACCTG 276
QY 252 TCTCCATTGAAGTTTCACTATGCTGTGCTGCTGCTGCTTCTTCTTCTTCTTCTTCT 311
Db 277 TCTCCATTGAAGTTTCACTATGCTGTGCTGCTGCTGCTTCTTCTTCTTCTTCTTCT 336
QY 337 ATTTTGGACATGTTTCACTATGCTGTGCTGCTGCTGCTTCTTCTTCTTCTTCTTCT 396
Db 372 ATGCTGGACATGTTTCACTATGCTGTGCTGCTGCTGCTTCTTCTTCTTCTTCTTCT 431
QY 397 ATGCTGGACATGTTTCACTATGCTGTGCTGCTGCTGCTTCTTCTTCTTCTTCTTCT 456
Db 432 GACTCCAGCTGTTTGGCATCAGGGGAGCTGATGGAACGTGCTTCTTCTTCTTCTTCTTCT 491
QY 457 GACTCCAGCTGTTTGGCATCAGGGGAGCTGATGGAACGTGCTTCTTCTTCTTCTTCTTCT 516
Db 492 TCATCAAAATTAATGATGTTGATGTTTAAAGATGCTCTTCTTCTTCTTCTTCTTCTTCT 551
QY 517 TCATCAAAATTAATGATGTTGATGTTTAAAGATGCTCTTCTTCTTCTTCTTCTTCTTCT 576
Db 552 TCTCTTAATGAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 611
QY 577 TCTCTTAATGAAAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 636
Db 612 GATPAAATGAGGTGCTGATGTTGAAAGCAATGATCTTGAATTTCTGCTGCTGAT 671
QY 637 GATPAAATGAGGTGCTGATGTTGAAAGCAATGATCTTGAATTTCTGCTGCTGAT 696
Db 672 TTTTCTTCAAGCAGCTTCTGATGAGAAAGAGTCTTCAAGTTTCTTCTTCTTCTTCTTCT 731
QY 697 TTTTCTTCAAGCAGCTTCTGATGAGAAAGAGTCTTCAAGTTTCTTCTTCTTCTTCTTCT 756
Db 732 TGTGTGAGAGATGCAAGTCAAAATTTGATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 784
QY 757 TGTGTGAGAGATGCAAGTCAAAATTTGATGTTTCTTCTTCTTCTTCTTCTTCTTCTTCT 816
Db 785 ----- 784
QY 817 GAATTAATAATPAAAGTACCTGATGGGACCTGTCTCTGTTCTGCTTGTGCTTTT 876
Db 785 ----- 784

Db 877 TCCCATGATGGGAGATGCTAGTCTCAGGGTCACTGAGTAAGTCTGTATATATATGAT 936
Qy 785 ----- 784
Db 937 ACTAATACTGAGAAATATACTTTCACACATGTGACACACCAGGTATGTACACTTGT 996
Qy 785 ----- 784
Db 997 GCTTTGCACTATATACCTTTTACTTGTCTATCTGTTCAATGACAAAACAGTGAATC 1056
Qy 785 -----AGCAAGGCGCAGACAACATCACTGAACCA 815
Db 1057 TGGCAATTTGACCTGAAAACATTTGCGACAGGCGCACAGAAACATCACTGAAGCAA 1116
Qy 816 TTTTACCAAGATTTGCTGAGAGAGAGTGTCTCAACATGCTTTGTGTGACAAAGATTAA 875
Db 1117 TTTTACCAAGATTTGCTGAGAGAGAGTGTCTCAACATGCTTTGTGTGACAAAGATTAA 1176
Qy 876 GATCTGTGGTATTTTCAAGATATATACATGATGGAAGAAAGACTGTTGAATCTTACA 935
Db 1177 GATCTGTGGTATTTTCAAGATATATACATGATGGAAGAAAGACTGTTGAATCTTACA 1236
Qy 936 AAAGAAAGTCTGCTGATGATTTGAAATTTG----- 966
Db 1237 AAAGAAAGTCTGCTGATGATTTGAAATTTGCTGAGATCTCTGCAATGCTGCTC 1296
Qy 967 -----AATCTTACGA 977
Db 1297 ACTGCACTCTCAACCTCTGCGGCTCAAGTATCTCTCTCACTCTGCAATCTCTGAGA 1356
Qy 978 CTGCGTGTGTAAGTCTGAGGAAATTTGAAGGCTCAGAGCAAGGTTAAATCCCTTCT 1037
Db 1357 CTGCGTGTGTAAGTCTGAGGAAATTTGAAGGCTCAGAGCAAGGTTAAATCCCTTCT 1416
Qy 1038 TCAGAAATCTCTGATGATTTATATATGTCATTAACCTAGAGAACTTATGAAGATCCGCTC 1097
Db 1417 TCAGAAATCTCTGATGATTTATATATGTCATTAACCTAGAGAACTTATGAAGATCCGCTC 1476
Qy 1098 ATGCTATCAGATGCTATTTCTATGAAAGAGACATGAAATTTGATGACAAAAG 1157
Db 1477 ATGCTATCAGATGCTATTTCTATGAAAGAGACATGAAATTTGATGACAAAAG 1536
Qy 1158 AAAGTCAAGTCTCCATGAGCAAAATCTTCTCTCAGCGGTAATTAACCAATAGG 1217
Db 1537 AAAGTCAAGTCTCCATGAGCAAAATCTTCTCTCAGCGGTAATTAACCAATAGG 1596
Qy 1218 ACTCTGAAGATGCGCATCATATAGATGCTGAGACACACAAAAGTAA 1266
Db 1597 ACTCTGAAGATGCGCATCATATAGATGCTGAGACACACAAAAGTAA 1645

RESULT 13

US-10-077-111-6
; Sequence 6, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, Jill
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA

ORGANISM: MOUSE
US-10-077-111-6

Query Match 36.5%; Score 464; DB 13; Length 1901;
Best Local Similarity 63.28; Pred. No. 2.5e-124;
Matches 916; Conservative 0; Mismatches 255; Indels 279; Gaps 2;

Qy 93 TTTCTTCAATTAAGAACATGATGTAAGTGAATTCACATTAAGTATGATGATGATGAT 152
Db 1 TTTCTTGTGTGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 60
Qy 153 GTCAATGCTGTGCTTCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 212
Db 61 GTCAATGCTGTGCTTCT 120
Qy 213 CGCTGTATCTGTTAGTGAATCTTACGAACTGCAATCTGCAATCTGCAATCTGCAAT 272
Db 121 CGCTGTATCTGTTAGTGAATCTTACGAACTGCAATCTGCAATCTGCAATCTGCAAT 180
Qy 273 TATGCTGTCAATGCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 332
Db 181 TATGCTGTCAATGCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Qy 333 GATGTAACAATGCTCTATGAAATCTGAATGACAGATGCTGCAATGATGAAACAG 392
Db 241 GATGTAACAATGCTCTATGAAATCTGAATGACAGATGCTGCAATGATGAAACAG 300
Qy 393 CTAATGTCAGCTCTGAGAGGTTTCCCAAGTCTCAGCTCTCAGCTCTCAGCTCTCAG 452
Db 301 CTAATGTCAGCTCTGAGAGGTTTCCCAAGTCTCAGCTCTCAGCTCTCAGCTCTCAG 360
Qy 453 GGGGAGCTGATGGAATGCTGTTTGTGGAATGACAGTATCAATATATATATATATAT 512
Db 361 GGGGAGCTGATGGAATGCTGTTTGTGGAATGACAGTATCAATATATATATATATAT 420
Qy 513 GGTAGTGTAAAGTGTCTCTGCGGCAATGCAATTTCTCTATGGAAGCTTCTT 572
Db 421 GGTAGTGTAAAGTGTCTCTGCGGCAATGCAATTTCTCTATGGAAGCTTCTT 480
Qy 573 GTCACTGCTCTCAGTGTGATTTAACAATGCTGGAATATTAATGATGATGATGAT 632
Db 481 GTCACTGCTCTCAGTGTGATTTAACAATGCTGGAATATTAATGATGATGATGAT 540
Qy 633 AGTGAAGAACATGATCTGGAATTAACCTGCTGCAATTTCTCTACAGCAATCTT 692
Db 541 AGTGAAGAACATGATCTGGAATTAACCTGCTGCAATTTCTCTACAGCAATCTT 600
Qy 693 GATGAGAACAGTCTTCAATTTTGAATGCACTGCAATGATGATGATGATGATGATGAT 752
Db 601 GATGAGAACAGTCTTCAATTTTGAATGCACTGCAATGATGATGATGATGATGATGAT 657
Qy 753 AAAATTTGATGTTTCTTTTACCATATCTT----- 784
Db 753 AAAATTTGATGTTTCTTTTACCATATCTT----- 784
Qy 718 CTAATGAGCACTGCGCCCTGTTCTGCGCTGCTGCTTTTCAATGATGAAAGATCTT 777
Db 718 CTAATGAGCACTGCGCCCTGTTCTGCGCTGCTGCTTTTCAATGATGAAAGATCTT 777
Qy 785 ----- 784
Db 785 ----- 784
Qy 778 GCATGGGGTCAATGATTAATCTGTCATATATATGATGATGATGATGATGATGATGAT 837
Db 778 GCATGGGGTCAATGATTAATCTGTCATATATATGATGATGATGATGATGATGATGAT 837
Qy 785 ----- 784
Db 785 ----- 784
Qy 838 CACAGCTGATCAGATACAGGATATGATGATGATGATGATGATGATGATGATGATGAT 897
Db 838 CACAGCTGATCAGATACAGGATATGATGATGATGATGATGATGATGATGATGATGAT 897
Qy 785 ----- 784
Db 785 ----- 784
Qy 898 TTAATGCTATGCTTCAATGACAAAGACATGATGATGATGATGATGATGATGATGATGAT 957
Db 898 TTAATGCTATGCTTCAATGACAAAGACATGATGATGATGATGATGATGATGATGATGAT 957
Qy 785 -----AGCAAGGCGCAGAAACATCACTGAGCAATTTACGGAAGATGGTCAAG 836
Db 785 -----AGCAAGGCGCAGAAACATCACTGAGCAATTTACGGAAGATGGTCAAG 836

Db 958 CCTTCCAGACAGAGCATGAAAGACCCCTGTAACATTCACTGAAGATGGTCAG 1017
QY 837 GAGTCGTCTCAACATGGCTTTGTGCACAGATTAAAGATCTTGTGTAATTTTCAG 896
Db 1018 GAGAGATCTCCCGTGGCTTCGCTCAGAGGCTTGAAGACCTCGTCGGTATTTTCAG 1077
QY 897 ATGATTAACATTGATGGAAGAAACCTGTGAATCTTACAAAAGAAAGTCTGGCTGAT 956
Db 1078 GCAACCAACATCATGGGAAGAACTATTCATCTCAAAAGGAAAGTCTGGCTGGAT 1137
QY 957 TTGAAAATTTGAATCTCTAGACCTCGTAAATGATGTAAGAAATTTGAAGCTCAG 1016
Db 1138 TTGAAAATTCGAATCTTAGGGCTGCGCAAGAAAGTCTGAGAGTATTTGAAGCTCAG 1197
QY 1017 ACCAAGTTAAATCCCTTTCTTCTGAGAAATCTGATGAATTTATGTCGAATTAAC 1076
Db 1198 GCCAAGATGATTCCTCTCTCTTCGGAATCCCTGACAGATTCATCTGCCAATTAACGA 1257
QY 1077 GAACTTATGAAAGATCCGATCATCGCATCAGATGGCTATTTCATATGAAAAGAACATG 1136
Db 1258 GAACTCATGAAAGACCCGTCATCGCATCAGATGGCTACTCTACGAGAGAGAACATG 1317
QY 1137 GAAAATTTGATTCAGCAAAAAGAAAGCTACAAAGTCCCATGACAAATCTTTCTCTTCA 1196
Db 1318 GAAAGCTGATTCACAAAGAAAGAGCTACAGACCCCATGACAAATTTGCTCTCTTCA 1377
QY 1197 GCGGTACTTACCAAAATAGACTGTGAAATGGCCATCAATTAATGATGGCTGAGACAC 1256
Db 1378 CTGTACTGACCCCAACAGACACTGAGATGGCCATCAACCGATGGCTGAGACGAC 1437
QY 1257 CAAAGTAA 1266
Db 1438 GAGAGTGAA 1447

RESULT 14
US-10-077-111-5
; Sequence 5, Application US/10077111
; Publication No. US20020187492A1
; GENERAL INFORMATION:
; APPLICANT: Todderud, C. Gordon
; APPLICANT: Finger, Joshua N.
; APPLICANT: Rillema, J111
; TITLE OF INVENTION: TBA
; FILE REFERENCE: 3053-4114US2
; CURRENT APPLICATION NUMBER: US/10/077,111
; CURRENT FILING DATE: 2002-02-15
; PRIOR APPLICATION NUMBER: 60/294,181
; PRIOR FILING DATE: 2001-05-29
; PRIOR APPLICATION NUMBER: 60/269,366
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 5
; LENGTH: 630
; TYPE: DNA
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: 630 bp partial nucleic acid sequence of human
; OTHER INFORMATION: RET16 cDNA
US-10-077-111-5

Query Match 29.7%; Score 377.2; DB 13; Length 630;
Best Local Similarity 97.9%; Pred. No. 3.1e-99;
Matches 382; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 774 ACCATATCTTAGCAGCGCAGACATCAGCTGAGCAATTTACGAGATTTGTC 833
Db 241 ACCTTTGCCAGGAGCGCAGACATCAGCTGAGCAATTTACGAGATTTGTC 300
QY 834 GAGAGGTCTCTCAACATGGCTTTGTGCACAAAGATTTAAAGATCTTGTGTAATTTTC 893
Db 301 GAGAGGTCTCTCAACATGGCTTTGTGCACAAAGATTTAAAGATCTTGTGTAATTTTC 360

QY 894 AAGATGAATTAACATTGATGGAAGAAAGCTTTGAATCTTACAAAAGAAAGTCTGGCTGAT 953
Db 361 AAGATGAATTAACATTGATGGAAGAAAGCTTTGAATCTTACAAAAGAAAGTCTGGCTGAT 420
QY 954 GATTGAAAATTTGAATCTCTAGACCTCGTAAATGATGTAAGAAATTTGAAGCTC 1013
Db 421 GATTGAAAATTTGAATCTCTAGACCTCGTAAATGATGTAAGAAATTTGAAGCTC 480
QY 1014 AGGACCAAGTTAAATCCCTTTCTTCTGAGAAATCTGATGAATTTATGTCGAATTAAC 1073
Db 481 AGGACCAAGTTAAATCCCTTTCTTCTGAGAAATCTGATGAATTTATGTCGAATTAAC 540
QY 1074 AGAAGCTTATGAAAGATCCGATCATCGATCAGATGAGTCTATTTCATATGAAAAGAA 1133
Db 541 AGAAGCTTATGAAAGATCCGATCATCGATCAGATGAGTCTATTTCATATGAAAAGAA 600
QY 1134 ATGAAAATTTGATTCAGCAAAAAGAAAGCT 1163
Db 601 ATGAAAATTTGATTCAGCAAAAAGAAAGCT 630

RESULT 15
US-10-357-930-5029/c
; Sequence 5029, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 5029
; LENGTH: 446
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-357-930-5029

Query Match 26.3%; Score 335; DB 18; Length 446;
Best Local Similarity 96.3%; Pred. No. 5.5e-87;
Matches 365; Conservative 0; Mismatches 10; Indels 4; Gaps 2;
QY 412 GGGTTTCCAGTTTCCCAAGCTCAAGTG---TTGGCATAGGGGAGCTGATGAA 468
Db 445 GGTTCACAGTTTTCCTCCAGACTCCCAAGCTGTTTGGCATCAGGGGAGCTGATGAA 386
QY 469 CTGTGTTTGTGGAATGACAGCTCATCAAAATTAATGATGATGTAAGTAAAGATG 528
Db 385 CTGTGTTTGTGGAATGACAGCTCATCAAAATTAATGATGATGTAAGTAAAGATG 326
QY 529 GCTCTTGGGCGGATGATTTTCTCTAATGAAAGCTTTGTCACTGAGCTCTCAT 588
Db 325 GCTCTTGGGCGGATGATTTTCTCTAATGAAAGCTTTGTCACTGAGCTCTCAT 266

```

Qy 589 GTGGTGAATTTAACAGTGTGGGATGATAAATGAGGTGCTGCTAGTGTG-AAAAAGCACAT 647
Db 265 GTGGTGAATTTAACAGTGTGGGATGATAAATGAGGTGCTGCTAGTGTG-AAAAAGCACAT 206
Qy 648 GATCTTGAATTTACCTGCTGGGATTTTCTTCACAGCCAGTTTCTGATGAGAACAAAGT 707
Db 205 GATCTTGAATTTACCTGCTGGGATTTTCTTCACAGCCAGTTTCTGATGAGAACAAAGT 146
Qy 708 CTTCAGTTTTTTGACATGCGCATCATGTGTCAGAGATTGCCAAGTCAAAATTTGATTGTT 767
Db 145 CTTCAGTTTTTTGACATGCGCATCATGTGTCAGAGATTGCCAAGTCAAAATTTGATTGTT 86
Qy 768 TCTTTACCCATATCTTAG 786
Db 85 TCTTTACCCATATCTTAG 67

```

Search completed: February 5, 2005, 08:35:28
 Job time : 768 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 5, 2005, 11:43:17 ; Search time 230 Seconds
(without alignments)
2731.871 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVLHTLADHGDVNCACF.....LTPNRLKMAINRLETHOK 384

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues
Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=x1h
-O=/cgn2_1/USPIO.spool/US10077111/runat_04022005_121053_6829/app_query.fasta_1.583
-DB=Issued_Patents_NA -OPMT=fastap -SUFFIX=p2n.rn1 -MINMATCH=0.1 -LOOPECI=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsu2 -TRANS=human40.cdt
-LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US10077111 @CGN_1.1.17 @runat_04022005_121053_6829 -NCPV=6 -ICPV=3
-NO MMAP -LARGEDUDRY -NEG SCORES=0 -MAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------------------------|
| 1 | 1978 | 96.6 | 1844 | 4 | US-09-620-312D-768 Sequence 768, App |
| 2 | 234.5 | 11.5 | 2359 | 1 | US-08-188-582-4 Sequence 4, Appl1 |
| 3 | 234.5 | 11.5 | 2359 | 1 | US-08-646-715-4 Sequence 4, Appl1 |
| 4 | 227 | 11.1 | 7028 | 4 | US-09-949-016-4194 Sequence 4194, Ap |
| 5 | 227 | 11.1 | 7028 | 4 | US-09-949-016-4195 Sequence 4195, Ap |
| 6 | 227 | 11.1 | 7042 | 3 | US-09-093-508-1 Sequence 1, Appl1 |
| 7 | 227 | 11.1 | 7042 | 3 | US-09-435-115-1 Sequence 1, Appl1 |
| 8 | 227 | 11.1 | 7042 | 3 | US-09-086-310-1 Sequence 1, Appl1 |
| 9 | 227 | 11.1 | 7042 | 4 | US-09-690-364-21 Sequence 21, Appl |
| 10 | 227 | 11.1 | 7042 | 4 | US-09-949-016-159 Sequence 159, Appl |
| 11 | 227 | 11.1 | 7075 | 3 | US-09-093-508-15 Sequence 15, Appl |
| 12 | 227 | 11.1 | 7075 | 3 | US-09-435-115-15 Sequence 15, Appl |

| | | | | | |
|----|-------|------|------|---|---------------------------------------|
| 13 | 226.5 | 11.1 | 2130 | 4 | US-09-620-312D-145 Sequence 145, App |
| 14 | 220.5 | 10.8 | 937 | 4 | US-10-101-464A-251 Sequence 251, Appl |
| 15 | 218 | 10.6 | 5152 | 3 | US-09-690-364-10 Sequence 10, Appl |
| 16 | 217.5 | 10.6 | 3747 | 3 | US-09-690-364-17 Sequence 17, Appl |
| 17 | 217.5 | 10.6 | 7157 | 4 | US-09-949-016-4192 Sequence 4192, Ap |
| 18 | 217.5 | 10.6 | 7157 | 4 | US-09-949-016-4193 Sequence 4193, Ap |
| 19 | 217.5 | 10.6 | 7171 | 4 | US-09-949-016-971 Sequence 971, Appl |
| 20 | 216 | 10.6 | 2152 | 1 | US-08-188-582-17 Sequence 17, Appl |
| 21 | 216 | 10.6 | 2152 | 1 | US-08-646-715-17 Sequence 17, Appl |
| 22 | 203.5 | 9.9 | 1246 | 3 | US-09-302-769-22 Sequence 22, Appl |
| 23 | 198.5 | 9.7 | 2166 | 3 | US-09-184-001-1 Sequence 1, Appl1 |
| 24 | 198.5 | 9.7 | 2558 | 3 | US-09-184-001-3 Sequence 3, Appl1 |
| 25 | 195 | 9.5 | 7886 | 2 | US-08-751-189-2 Sequence 2, Appl1 |
| 26 | 195 | 9.5 | 7886 | 2 | US-09-060-836-2 Sequence 2, Appl1 |
| 27 | 195 | 9.5 | 7886 | 2 | US-09-184-445-2 Sequence 2, Appl1 |
| 28 | 193.5 | 9.5 | 2369 | 3 | US-09-302-769-20 Sequence 20, Appl |
| 29 | 188 | 9.2 | 3465 | 3 | US-08-914-999-5 Sequence 5, Appl1 |
| 30 | 187.5 | 9.2 | 9991 | 4 | US-09-902-540-1014 Sequence 1014, Ap |
| 31 | 187 | 9.1 | 1422 | 4 | US-09-248-796A-4567 Sequence 4567, Ap |
| 32 | 184.5 | 9.0 | 1157 | 4 | US-09-270-767-14727 Sequence 14727, A |
| 33 | 184.5 | 9.0 | 1731 | 4 | US-09-902-540-1920 Sequence 1920, Ap |
| 34 | 183.5 | 9.0 | 1548 | 4 | US-09-614-221A-136 Sequence 136, Appl |
| 35 | 182 | 8.9 | 2481 | 3 | US-08-899-578-1 Sequence 1, Appl1 |
| 36 | 180.5 | 8.8 | 1638 | 4 | US-09-902-540-8422 Sequence 8422, Ap |
| 37 | 180.5 | 8.8 | 6553 | 4 | US-09-902-540-885 Sequence 885, App |
| 38 | 180 | 8.8 | 1182 | 4 | US-09-248-796A-4568 Sequence 4568, Ap |
| 39 | 178.5 | 8.7 | 1542 | 4 | US-09-949-016-1365 Sequence 1365, Ap |
| 40 | 176 | 8.6 | 1185 | 4 | US-09-248-796A-4600 Sequence 4600, Ap |
| 41 | 176 | 8.6 | 2272 | 3 | US-09-108-857-1 Sequence 1, Appl1 |
| 42 | 175.5 | 8.6 | 1115 | 1 | US-08-190-802A-19 Sequence 19, Appl |
| 43 | 175.5 | 8.6 | 1115 | 3 | US-08-477-346-19 Sequence 19, Appl |
| 44 | 175.5 | 8.6 | 1115 | 3 | US-08-473-089-19 Sequence 19, Appl |
| 45 | 175.5 | 8.6 | 1115 | 3 | US-08-487-072A-19 Sequence 19, Appl |

ALIGNMENTS

RESULT 1
US-09-620-312D-768
Sequence 768 Application US/09620312D
Patent No. 656962
GENERAL INFORMATION:
APPLICANT: Tassadit, Tom
APPLICANT: Liu, Chenghua
APPLICANT: Neundi, Vinod
APPLICANT: Zhang, Yie
APPLICANT: Ren, Feiyun
APPLICANT: Chen, Rui-hong
APPLICANT: Zhao, Qing A.
APPLICANT: Wehrman, Tom
APPLICANT: Xue, Aйдong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yundong
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 656962el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP28
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: pt_Fl_Genes Version 1.0
SEQ ID NO 768
LENGTH: 1844
TYPE: DNA


```

: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1989
: TELEFAX: (415) 398-3249
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2359 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 49..2160
: US-08-188-582-4

Alignment Scores:
Pred. No.: 1,21e-17 Length: 2359
Score: 234.50 Matches: 88
Percent Similarity: 44.34% Conservative: 49
Best Local Similarity: 28.48% Mismatches: 128
Query Match: 11.46% Indels: 45
DB: 1 Gaps: 11

US-10-077-111-13 (1-384) x US-08-188-582-4 (1-2359)
Qy 3 LysleuIIeHstHrLeuAlaAspHsGlyAspAspValaEnCySyalAphSer--- 21
Db 1369 GAGGAAACAGAGAGCTTAATGGGTCAACCGGACCCGATACCGGTGGCTTGGCCCC 1428
Qy 22 --PheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTySeriLeuArg 40
Db 1429 GAGATGAACCTGTGGCTCTCATGTCCGAGGACAGCACCATPAGAGCTGTGCTGTGCTC 1488
Qy 41 AspHethrGluLeuProHisSerProLeuLysPheHstHrTyAlaValHisCyS 60
Db 1489 ACCGTGCTCTCGTA-----GTACCTACCGCGGGGACGTTTACCCGGTGGAGATGT 1542
Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
Db 1543 CGCTTGGCGGACATGGCTCATATTGTTCTTCTTGTCTGACGACAAACCTGCTGTG 1602
Qy 81 TrpAsnHrGluAsnGlyIleMerLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 1603 TGGGCGACGAGATTCACATCAACAGCGTTGGCGGATTCGAGGACACTTG--TCGGACG 1658
Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
Db 1660 GATTGTGTACAAATTCATCCCAATTCCAATATATGTGGCCACCGGTTCTAGCGATGCACG 1719
Qy 121 ValValLeuThrPAsnAlaGlnSerTyrlLysLeuTyArgCysGlySerValLysAspGly 140
Db 1720 GTACGCCCTGTGGAGAACATGACCCGCTCATGCGTACCGCTCATGATACGGGCCACAAGGA 1775
Qy 141 SerLeuAlaAlaCysAlaAphSerProAsnGlySerPhePheValThrGlySerSerCys 160
Db 1780 TCGGGAAGTTCCTGTGGCTTCTCGCCGCGGCGGATATCTGGGCTCGGGTTCACTAAT 1833
Qy 161 GlyAspLeuThrValTrpAsp---AspLysMetArgCysLeuHisSerGlyLysAlaHis 179
Db 1840 CACATCATCATCATCTGGGATCTGTGCAACGGATCCCTGTGCACACCCCTGTGAGGAC 1899
Qy 180 AspLeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGly 199
Db 1900 ACTAGCACTGTGACCAACGATCACTTTAGT-----CGGATGTGAACAGTC--- 1944
Qy 200 LeuGlnPheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleVal 219
Db 1945 -----CTGGCTCGACCGGCTTGATTAACAATCTAATCTGTGG----- 1983
Qy 220 SerPheThrHisIleLeuAlaArgArgThrGluHisGlnLeuLysGlnPheThrGluAsp 239

```

```

Db      1984  -----GACITTCACAAGGTTACGGAAGC 2007
QY      240  TTPSerGIuGIuValISerThrTTPLeuCyelAGIuAspleuLyAspleuValGIly 259
      :::::|||||
Db      2008  TATATCAGCATCATCATC---ACTGTGTCCACCATCAGATGAGAAAGC---GAGCAGC 2061
QY      260  ILeheLySeMetAsnAnuIIeAspGIlyLeuGIu-----LeuLeuAsnLeuThr 275
      :::::|||||
Db      2062  GTTCACTCACTGAGGCTACTTTCCCGACAGAAACTGCCATTGTGTACGCTGCACTTTAGC 2121
QY      276  IyGIuSerLeuAlaAsp-----AspleuLySIIeGIuSer 287
      :::::|||||
Db      2122  CGCGGAATATCTCTT-CATGTGCGGSGTCTATTTCAGAGTTAGAGACAGATAGCTTA 2180
QY      288  LeuGIlyLeuArgSerLySValIleuArg 296
      :::::|||||
Db      2181  TTTGGTATACGTAATGATGATTAGG 2207

RESULT 3
US-08-646-715-4
: Sequence 4, Application US/08646715
: Patent No. 5637686
: GENERAL INFORMATION:
: APPLICANT: Tjian, Robert
: APPLICANT: Comati, Lucio
: APPLICANT: DynIact, Brian D.
: APPLICANT: Hoey, Timothy
: APPLICANT: Ruppert, Siegfried
: APPLICANT: Tanese, Naoko
: APPLICANT: Wang, Edith
: APPLICANT: Weinzierl, Robert O.J.
: TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
: TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
: NUMBER OF SEQUENCES: 36
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: FLEHR, HOEBACH, TEST, ALARITTON & HERBERT
: STREET: 4 Embarcadero Center, Suite 3400
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-4187
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/646,715
: FILING DATE: 09-MAY-1996
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/188,582
: FILING DATE: 28-JAN-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Osman, Richard A
: REGISTRATION NUMBER: 36,627
: REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 781-1869
: TELEFAX: (415) 398-3249
: TELEX: 910 277299
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2359 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 49..2160
: US-08-646-715-4

```



```
Qy 163 LeuThrVal1TrpAspArgCysLeuHisSer----- 175
    :::::|||||
    3131 ATCAGGCTCTGGAGACAAAGAAATGTATGTAAGAACTGTGTATGTTAAAGCAAGAA 3190
Qy 175 ----- 175
    3191 GTAGATGTGTGTTTCAAGAAATGAATGATGTCCTTGACATGACATATAGACGT 3250
Qy 176 -----GluTyrSalHisAspLeuGlyIleThr 184
    3251 CTGCACTCATTTATGAGAACAGGTCAAGTATGATATCTGACTGAAGCTCAAGTTAGC 3310
Qy 185 CysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlyLeuGlnPhePheArg 204
    3311 TGCGTGTGCTTAAGTCCACAT-----CTTCAGTAC----- 3340
Qy 205 LeuAlaSerCysGlyGlnAspCysGlnValIleTrpIleValSerPheThrHisIle 224
    3341 ATTCGATTTGGAGATGAATAATGAGACCATGATTTAGAACTGTGAAACAATAGATC 3400
Qy 225 LeuAlaArgArgThrGlnHisGlnLeuLys-----GlnPheThrGlnAspTrp 240
    3401 TTCGATTCAGGTTTCAGACACAGAAACCTGATGACATCCAGTTCACAGCCGAT-- 3457
Qy 241 SerGlnGluValIleSerThrTrpLeuCysAlaGlnAsp-LeuLysAspLeuValGlyIle 260
    3458 GAGAAAGCTCTTATTTCAAGT-----TCTGATGATGTGTAATTCAGATATGAAAT 3508
Qy 260 ePheLysMetAsnAsnIleAspGlyLysGlnLeuLysLeuThrLysGlnSerLeuAla 280
    3509 TGGCAATTTGACAAATGATATCTTCTACGAGGCCATCAGAAACAGTGAAGACTTAA-- 3566
Qy 280 AspAspLeuLysIleGlnSerLeuGlyLeu-----ArgSerLys 293
    3567 -GACTCTTGAATAATTCAGACCTGCTTCTTGTCATTTGATGACAGACAGTAAAGTATGC 3625
Qy 293 eValLeuArgLysIleGlnLeuArgThrLysValLysSerLeuSer----- 309
    3626 AATATTAA-----TTACTGGAATAAAGAAAGAAACATTTGCTCTCCACAGGATACATA 3679
Qy 310 -----SerGlyIleProAspGlnPheIleCysProIleThrArgI 323
    3680 CTTTCTTGATGACATTTCTCAGATGCTACCA--AGTTTCATCTACTGCTGACAGAG 3736
Qy 323 uLeuMetLysAspProValIleAlaSer 332
    3737 ACTGCAAGATCTGAGTTTGTATCTCC 3764
Db
RESULT 5
US-09-949-016-4195
; Sequence 4195, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 4195
; LENGTH: 7028
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-4195
```

Alignment Scores:

| | | | |
|------------------------|----------|---------------|------|
| Pred. No.: | 6,63e-16 | Length: | 7028 |
| Score: | 227.00 | Matches: | 102 |
| Percent Similarity: | 40.24% | Conservative: | 63 |
| Best Local Similarity: | 24.88% | Mismatches: | 140 |
| Query Match: | 11.09% | Indels: | 105 |
| DB: | 4 | Gaps: | 19 |

US-10-077-111-13 (1-384) x US-09-949-016-4195 (1-7028)

```
Qy 3 LysLeuIleHisThrLeuAlaAspHisGlyAspArgValLysCysValAlaPheSerPhe 22
    :::::|||||
    2612 GAATGATGACACCTATGATGAGCACTCAGACCAATGCTTCCATTCACCAAC 2671
Qy 23 Ser-----LeuAlaIlePheCysSerLeuAspLysThrIleArgLysSer 38
    2672 AGTAGTCATGATCTTCTTACGACCTGGGTCAAGTACGCTTCCCAAACTTTGGGAT 2731
Qy 39 LeuArgAspPheThrGlnLeuProHisSerProLeuLysPheHisThrTrpAlaValHis 58
    2732 TTGAATGAA-----AAGAAATGTCGAAATACCATGTTGTGTCATACAAATTCAGTCAAT 2785
Qy 59 CysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThr 78
    2786 CACTGCAGATTTTACACAGATGATTAAGCTTTGGCTAGTTGTTCCAGCTGAGAACCTTA 2845
Qy 79 ValLeuTrpAsnThrGlnGlnGly-----GlnMetLeu 89
    2846 AAGCTTTGGATGCGCATCAGCAAAATGAGAGAAAGCAATTAATGTAACAGTTCTTC 2905
Qy 90 AlaValMetGlnGlnProSerGlySerPro-----ValArgValCysGlnPheSer 106
    2906 CTAAATTTGGAGACACCTCAAGAGATATGGAATGATATGTAAGCTTTGCTGCTCT 2965
Qy 107 ProAspSerThrCysLeuAlaSerGlyValAlaIleAspGlyThrValValLeuTrpAsnAla 126
    2966 GCTGATGTCGACAGATATGATGTTG--GCAGCAAAAATAAATCTTTTGTGGAAATACA 3022
Qy 127 GlnSer---TyrLysLeuLysTrpArgCysGlySerValLysAspGlySerLeuAla----- 143
    3023 GACTCAGGTTCAAGAGGCTGATGTC-----AGAGACACTTTTAAGTTGGGTT 3070
Qy 144 ---AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAsp 162
    3071 CATGCTGATGATGTTTCTCTGATGATCATCTTTTGACATCTTCTGATGACCAACA 3130
Qy 163 LeuThrVal1TrpAspArgCysLeuHisSer----- 175
    3131 ATCAGGCTCTGGAGACAAAGAAATGTATGTAAGAACTCTGTAATGTTAAAGCAAGAA 3190
Qy 175 ----- 175
    3191 GTAGATGTGTGTTTCAAGAAATGAATGATGTCCTTGACATGACATATAGACGT 3250
Qy 176 -----GluTyrSalHisAspLeuGlyIleThr 184
    3251 CTGCACTCATTTATGAGAACAGGTCAAGTATGATATCTGACTGAAGCTCAAGTTAGC 3310
Qy 185 CysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlyLeuGlnPhePheArg 204
    3311 TGCGTGTGCTTAAGTCCACAT-----CTTCAGTAC----- 3340
Qy 205 LeuAlaSerCysGlyGlnAspCysGlnValIleTrpIleValSerPheThrHisIle 224
    3341 ATTCGATTTGGAGATGAATAATGAGACCATGATTTAGAACTGTGAAACAATAGATC 3400
Qy 225 LeuAlaArgArgThrGlnHisGlnLeuLys-----GlnPheThrGlnAspTrp 240
    3401 TTCGATTCAGGTTTCAGACACAGAAACCTGATGACATCCAGTTCACAGCCGAT-- 3457
Qy 241 SerGlnGluValIleSerThrTrpLeuCysAlaGlnAsp-LeuLysAspLeuValGlyIle 260
    3458 GAGAAAGCTCTTATTTCAAGT-----TCTGATGATGCTGAAATTCAGGATATGAAAT 3508
Db
```



```

Db      3509 TGGCAATTGACAAATGATCTTTCTACGAGGCCATCAGAAACATGAAAGACTTAA-- 3566
Qy      280 AASPAPLeuLysIleGluSerLeuGlyLeu-----ArgSerly 293
Db      3567 -GACTCTTGAATAATTCAGACCTGCTTTCTTGTCATTTGATGGAACAGTGAAGTATGC 3625
Qy      293 sValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeuSer----- 309
Db      3626 AATATTAA-----TTACTGGAATAAAGAAAGAAACCTTGTCTGCACAGGCTACAGTA 3679
Qy      310 -----SerGlyIleProAspGluPheIleCysProIleThrArgI 323
Db      3680 CTTTCTTGACATTTCTACAGATGCTACCA---AGTTTCATCTACTCTGCTGACAG 3736
Qy      323 uLeuMetLysAspProValIleAlaSer 332
Db      3737 ACTGCAAAAGATCTGAGCTTTGATCTCC 3764

RESULT 7
US-09-435-115-1
; Sequence 1, Application US/09435115
; Patent No. 6346607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF C ASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 6346607 West Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435,115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USU1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7042 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578...4159
; OTHER INFORMATION:
; US-09-435-115-1

Alignment Scores:
Pred. No.: 6,65e-16 Length: 7042
Score: 227.00 Matches: 102
Percent Similarity: 40.24% Conservative: 63

```

```

Best Local Similarity: 24.88% Mismatches: 140
Query Match: 11.09% Indels: 105
DB: 3 Gaps: 19

US-10-077-111-13 (1-384) x US-09-435-115-1 (1-7042)
Qy      3 LysLeuIleHisThrLeuAlaAspHisGlyAspArgValIleAspCysAlaPheSerPhe 22
Db      2612 GAAGTGTACACACCTATGATGAGCACTCAGACAGATGATGCTGCTGCTTCCACCAAC 2671
Qy      23 Ser-----LeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSer 38
Db      2672 AGTAGTCATCATCTTCTCTTACGACCTGCGCTCAAGTCACTGCTCCCAACTTGGGAT 2731
Qy      39 LeuArgAspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHis 58
Db      2732 TTGAATCAA-----AAAGAAATGTCGAATATCAATGTTGCTCATACAAATTCAGTCAAT 2785
Qy      59 CysCysAspSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThr 78
Db      2786 CACTGCAGATTTTCCACAGATGATAGCTTTGGCTAGTTGTTACGCTGATGGAACCTTA 2845
Qy      79 ValLeuTrpAsnThrGluAsnGly-----GlnMetLeu 89
Db      2846 AAGCTTTGGGATGCGACATCAGCAAAATGAGAGGAAACATTAATGTAACAGTTCTTC 2905
Qy      90 AlaValMetGluGlnProSerGlySerPro-----ValArgValCysGlnPheSer 106
Db      2906 CTAAATTTGAGAGACCTCAAGAGATGAGAAAGTATGATAGTGTGTGCTGCTCT 2965
Qy      107 ProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValIleuTrpAsnAla 126
Db      2966 GCTGATGCTGCAGATTAATGCTG---GCGACAAAAAATAAATCTTTTGTGTGAATACA 3022
Qy      127 GlnSer---TyrLysLeuTyrArgCysGlySerValLysAspGlySerLeuAla----- 143
Db      3023 GACTCAGCTTCAAAAGGCGCTGATGCTC-----AGAGAGCATTTAAAGTTGGGCTT 3070
Qy      144 ---AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAsp 162
Db      3071 CATGCTGTGATGTTTCTCTGATGATCATATTTTGAACATCTTGTGATGACACAGACA 3130
Qy      163 LeuThrValTrpAspAspLysMetArgCysLeuHisSer----- 175
Db      3131 ATCAGGCTCTGGAGACAAAGAAAGATGTAAGAATCTCTGTAATGTTAAAGCAAGAA 3190
Qy      175 ----- 175
Db      3191 GTAGATGTTGTTTCAAGAAATGAAGTATGATGCTCTGCACTGACCATATTAAGACGT 3250
Qy      176 -----GlnLysAlaHisAspLeuGlyIleThr 184
Db      3251 CTGCAACTCATTAATGGAAGACAGTCAATGATTAATGCTGACAGTCAAGTTAGC 3310
Qy      185 CysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeuGlnPhePheArg 204
Db      3311 TGCTGTGCTTAAGTCCACAT-----CTTCAGTAC----- 3340
Qy      205 LeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSerPheThrHisIle 224
Db      3341 ATTGCATTTGAGATGATGAATGAGGCATGATTAATTTAGAACTGTGAACCAATAGATC 3400
Qy      225 LeuAlaArgArgThrGluHisGlnLeuLys-----GlnPheThrGluAspTrp 240
Db      3401 TTCAGTCCAGGTTTTCAGCAAGAAACCTGATGCGACATCCAGTTCACAGCCGAA--- 3457
Qy      241 SerGluGluValValSerThrTrpLeuCysAlaGlnAsp-LeuLysAspLeuValGlyI 260
Db      3458 GAGAGACTCTTATTTTCAGT-----TTGATGATGCTGAAATTCAGTATGGAAT 3508
Qy      260 ePheLysMetAsnAsnIleAspGlyLysGluLeuAsnLeuThrLysGluSerLeuAl 280
Db      3509 TGGCAATTGACAAATGATCTTTCTACGAGGCCATCAGAAACATGAAAGACTTAA-- 3566

```

[illegible]

Oy 107 ProaSerThrCysLeuAlaSerGlyAlaAlaAAspGlyThrValValLeuTProAsnAla 126
 Db 2366 GCTGATGCTGCACAGGATATGTCG--GCAGCAAAAAATAAATCTTTTGTGTGAATACA 3022
 Oy 127 GlnSer---TyrIySLeuTyraGcYsGlySerValIyAspGlySerLeuAla----- 143
 Db 3023 GACTCAAGCTTCAAAAGATGCGCTGATTCG-----AGAGGACATTTAAAGTGGGCTT 3070
 Oy 144 ---AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAsp 162
 Db 3071 CAGGTGTGATGTTTCTCTGTATGATCACTCATTTTGTGACATCTTGTGATGACAGACA 3130
 Oy 163 LeuThrValItrpAspAspLeuAspCysLeuHisSer----- 175
 Db 3131 ATCAGGCTCTGGAGACAAAGAAAGTATGTAAGAATCTCTGTATGTTAAAGCAAGAA 3190
 Oy 175 ----- 175
 Db 3191 GTAGATGTGTGTTTCAAGAAATGAAGTAGATGGTCTTCGATGACATATTAAGCGT 3250
 Oy 176 -----GluIySAlaHisAspLeuGlyIleThr 184
 Db 3251 CTCGAATCTCATTAATGAGAAACAGATCAGATTTGATATCTGACTGAAGCTCAAGTAGAC 3310
 Oy 185 CysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeuGlnPhePheArg 204
 Db 3311 TGCCTGTGCTTAAGTCCACAT-----CTTCAGATAC----- 3340
 Oy 205 LeuAlaSerCysGlyGlnAspCysGlnValIyLeuTrpIleValSerPheThrHisIle 224
 Db 3341 ATTGCATTTGGAGATGAAATATGAGCGCATTTGAGATTTTAACTGTAAACATATGATC 3400
 Oy 225 LeuAlaArgArgThrGluHisGlnLeuLys-----GlnPheThrGluAspTrp 240
 Db 3401 TTCACATGCCAGGTTTCGCACAGAAACCTGATGCGACATCCAGTTCACAGCCGAT-- 3457
 Oy 241 SerGluGluValIyAlaSerThrTrpLeuCysAlaGlnAsp-LeuIyAspLeuValGlyI 260
 Db 3458 GAGAAAGACTCTTATTTACAGT-----TCTGATATGCTGAATATTCAGATGAGAAAT 3508
 Oy 260 ePheIyMeCAsnAsnIleAspGlyIyGluLeuLeuAsnLeuThrIyGluSerLeuAl 280
 Db 3509 TGGCAATTTGGACAAATATATCTTTCTACAGGCGCATGAGAAACAGTGAAGACTTAA-- 3566
 Oy 280 AAspAspLeuIyIleGluSerLeuGlyLeu-----ArgSerIy 293
 Db 3567 -GACTCTTGA AAAATTCAGAGACTGCTTCTTGTCATTTGATGGAACAGTGAAGGTATGC 3625
 Oy 293 eValLeuArgIyIleGluGluLeuAspThrIyValIySerLeuSer----- 309
 Db 3626 AATATTTA-----TTACTGAAAATAAGAAAGAAAGACTTTGTCTGTCACAGGGTACAGTA 3679
 Oy 310 -----SerGlyIleProAspGluPheIleCysAspProIleThrArgI 323
 Db 3680 CTTTCTGTGACATTTCTACAGATGCTACCA--AGTTTCATCTACTCTCTGCTGACAG 3736
 Oy 323 uLeuMeIyAspProValIleAlaSer 332
 Db 3737 ACTGCAAAAGATCTGAGACTTTGATCTCC 3764
 RESULT 9
 ; US-09-690-364-21
 ; Sequence 21, Application US/09690364
 ; Patent No. 6468795
 ; GENERAL INFORMATION:
 ; APPLICANT: Hong Zhang
 ; APPLICANT: Andrew T. Walt
 ; TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
 ; FILE REFERENCE: RTS-0190
 ; CURRENT APPLICATION NUMBER: US/09/690,364
 ; CURRENT FILING DATE: 2000-10-17
 ; NUMBER OF SEQ ID NOS: 100
 ; SEQ ID NO 21

| | | | | |
|----|--|------|--|------|
| Oy | | 59 | CysCyCysPheSerProSerGlyHisIleLeuMetLaserCysSerThrAspGlyThrThr | 78 |
| Dd | | 2819 | CACTCGCAGATTTCACCCAGATGAATTAAGCTTTGGCTAAGTGTTCACGCTGATGAAACTTA | 2878 |
| Oy | | 79 | ValLeuTPraenthrgluAsngLy-----GlnMetLeu | 89 |
| Dd | | 2879 | AAGCTTTGGATGGCACTTCAGACAATAAGAGAAGAAAACATTAAATGTGAACCATGCTTTC | 2938 |
| Oy | | 90 | AlaValMetGluGlnProSerGlySerPro-----ValArgValCysGlnPheSer | 106 |
| Dd | | 2939 | CTAAATTGTGGAGACCCTCCAAGAGATATGGAAGTAGATAGTGAAGTGTTGTTCGGTGCT | 2998 |
| Oy | | 107 | ProAspSerThrCysLeuAlaSerGlyValAlaIleAspGlyThnValValLeuTriPAsnAla | 126 |
| Dd | | 2999 | GCTGATGCTGCAGAGATTAATGCTG---CCACCAAAAAATTAATCTTTTGTGGAAATCA | 3055 |
| Oy | | 127 | GlnSer---TyrlsLeuTyArGcYasGlyserVallysAspGlySerLeuAla----- | 143 |
| Dd | | 3056 | GACTCACCGTTCAAAGTGCGTCAATTGC-----AAGAGCATTTAAGTTCGGGCT | 3103 |
| Oy | | 144 | ---AlaCysAlaPheSerProAsngLySerPhePheValTrnGlySerSerCysGlyAsp | 162 |
| Dd | | 3104 | CATGCTGTGATGATTTCCTCGATGATCAATCAATTTTGACATCTTCGTGATGACCAACA | 3163 |
| Oy | | 163 | LeuThrValTrlPaaspAspLysMeArGcYsLeuHisSer----- | 175 |
| Dd | | 3164 | ATCAGCGCTCTGGAGACAACAAGAAAGTATGTAAAGAACCTCGCTGTAAATGTTAAGCAAGA | 3223 |
| Oy | | 175 | ----- | 175 |
| Dd | | 3224 | GTAAGATGTGTTTTCAAGAAAATGMACTGATGCTCTTGCAAGTTGACATATAGAAGCT | 3283 |
| Oy | | 176 | -----GluylsAlaHisAspLeuGlylIethr | 184 |
| Dd | | 3284 | CTGCAACTCATTTAATGAGAAGAAGCAAGCTCAATTAATTCGACTGAAGCTCAAGTTAGC | 3343 |
| Oy | | 185 | CysCyAspPheSerSerGlnProValSerAspGlyGlnGlyLeuGlnPhePheArg | 204 |
| Dd | | 3344 | TGCTGTTCCTTAAGTCCACAT-----CTTCAGATAC----- | 3373 |
| Oy | | 205 | LeuAlaSerCysGlyGlnAspCyselInValysIleTrpIleValSerPheThrHisIle | 224 |
| Dd | | 3374 | ATTGATTTTGGAGATGAATAATGAGCCCATGGAATTTAAGACTTGTAACAATATGAATC | 3433 |
| Oy | | 225 | LeuAlaArgTrhGlnHisGlnLeuLys-----GlnPheThrGlnAspTrp | 240 |
| Dd | | 3434 | TTCCAGCTCCAGGTTTCACACACAGAAAATCTGATGGCACATCCAGCTTCACAGCCGAT--- | 3490 |
| Oy | | 241 | SerGlnGluValValSerThrTrlPLeuCysalaGlnAsp-LeuLysAspLeuValGlylI | 260 |
| Dd | | 3491 | GAGAAAGACTTATATTTCAAGT-----TCTGATGATGCTGAATTCAGGTATGAGAT | 3541 |
| Oy | | 260 | ePheLysMeArAsnAniLeaspGlyLysGlnLeuLeuAsnLeuThrlLysGlnSerLeuAl | 280 |
| Dd | | 3542 | TGGCAATTGGACAATATATCTTTTACAGAGGCCATTCAGAAACAGTGAAGAAACATTTA- | 3599 |
| Oy | | 280 | aAspAspLeuLysIleGlnSerLeuGlyLeu-----ArgSerLy | 293 |
| Dd | | 3600 | -GACTCTTGAAAAATTCAAAGACTGCTTTCGTGCATTTGATGGAACAGTAAGGTATGG | 3658 |
| Oy | | 293 | sValLeuArgLysIleGlnGluLeuArgThrLysValLysSerLeuSer----- | 309 |
| Dd | | 3659 | AATATTA-----TTACTGGAATATAAAGAAAAGCTTGTCGTGCACAGGGTACAGTA | 3712 |
| Oy | | 310 | -----SerGlyIleProAspGlnPheHisLecysProIleThrArgI | 323 |
| Dd | | 3713 | CTTCTTGATGACATTTTCACAGATGCTNACA---AGTTTATATCAACTCGTGCAGAAAG | 3766 |
| Oy | | 323 | uLeuMetLysAspProValIleAlaSer | 332 |
| Dd | | 3770 | ACTGCAAGATCTGGAGTTTATGATCTCC | 3797 |

```

RESULT 12
US-09-435-115-15
; Sequence.15, Application US/09435115
; Patent No. 634607
; GENERAL INFORMATION:
; APPLICANT: Henzel, William J.
; TITLE OF INVENTION: APAF-1, AN ACTIVATOR OF CASPASE-3
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant, Gould, Smith, Edell, Welter & Schmidt
; STREET: 3100 No. 634607west Center, 90 South Seventh St
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/435.115
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/092,508
; FILING DATE:
; APPLICATION NUMBER: 60/055,258
; FILING DATE: 07-AUG-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Kettelberger, Ph.D., Denise M
; REGISTRATION NUMBER: 33,924
; REFERENCE/DOCKET NUMBER: 11669.6USUI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; TEX:
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7075 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 578..4192
; OTHER INFORMATION:
; US-09-435-115-15

Alignment Scores:
Pred. No.: 6.7e-16 Length: 7075
Score: 227.00 Matches: 102
Percent Similarity: 40.24% Conservative: 63
Best Local Similarity: 24.88% Mismatches: 140
Query Match: 11.09% Indels: 105
DB: 3 Gaps: 19

US-10-077-111-13 (1-384) x US-09-435-115-15 (1-7075)

QY 3 LysLeuLeuLeuHisThrLeuLeuAlaAspHisGlyValAspAspValAlaGlyCysValAlaPheSerPhe 22
Db 2645 GAACCTAGTACACACACTGATGATGACGACTGAGCAGCAAGCAATTCCTGTCATTCACCAAC 2704
QY 23 Ser-----LeuLeuAlaThrCysSerLeuAspArgLysThrIleArgLeuTyrSer 38
Db 2705 AGTAGTCATCATCTTCTCTTAGCCACTGGGCTCAAGTACGCTTCCTCAACATTTGGGAT 2764
QY 39 LeuAlaGAspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHis 58
Db 2765 TTGAATCA-----AAAGAATGCGAAATATCCATGTTGGTCTATACAAATCTAGTCAAT 2818
QY 59 CysCysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThr 78

```

Db 2819 CACTGAGATTTTACCAGATGATTAAGCTTTGGCTAGTGTTCAGCTGATGACACCTTA 2878
79 ValLeuTriPheThrGluLeuGly-----GlnMetLeu 89
Db 2879 AAGCTTTGGATGCGCATGACCAATGAGAGAAAAGCATTAATGGAACAGTTCTTC 2938
90 AlaValMetGluInProSerGlySerPro-----ValArgValCysGlnPheSer 106
Db 2939 CTAAATTTGGAGACCTTCAAGAGATATGGAAGTATGTAAGTGTGTGTGTGTGTGTGT 2998
107 ProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThrValValLeuTriPheAla 126
2999 GCTGATGCTGCAAGATTAATGCTG--GCAGCAAAAATAATTAATCTTTTGTGGAATACA 3055
127 GlnSer---TyrLeuSerLeuTyrArgCysGlySerValIleAspGlySerLeuAla----- 143
3056 GACTCAGCTTCAAAAGGCGCTGATGC-----AGAGACATTTTAAGTTGGGTT 3103
144 ---AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAsp 162
3104 CATGTGTGATGTTTCTCTCTGATGATCATCAATTTTGTGACATCTTCTGATGACAGACA 3163
163 LeuThrValIleTrpAspIlePheAspIleMetArgCysLeuHisSer----- 175
3164 ATCAGGCTCTGGAGACAAAGAAAGATATGTAGAACCTGCTGTAATGTTAAGCAAGAA 3223
175 ----- 175
3224 GTAGATGTTGTCTTTCAGAAAATGAAGTATGCTGCTTCAGTGCATATTAAGACGT 3283
176 -----GluLeuAlaHisAspLeuGlyIleThr 184
3284 CTCGAATCATTAATGAGAAAGACAGGTCAAGTATGATATCTGACTGAAGCTCAAGTTAGC 3343
185 CysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlnPhePheArg 204
3344 TGCTGTGCTTAAGTCCACAT-----CTTCAGTAC----- 3373
205 LeuAlaSerCysGlyGlnAspCysGlnValIleTyrIleValSerPheThrHisIle 224
3374 ATTGCATTTGGAGATGAAAATGAGACCATTAAGATTTTGAACCTTGTAAACATAGACATC 3433
225 LeuAlaArgArgThrGlnHisGlnLeuIle-----GlnPheThrGluAspTyr 240
3434 TTCAGATCCAGGTTTTCAGCAAGAAAGATGATGACATCCAGTTCACAGCCGAT-- 3490
241 SerGluGluValIleSerThrTyrPheCysAlaGlnAsp--LeuIleAspLeuValGlyIle 260
3491 GAGAAACCTCTTAATTCACAT-----TCTGATGATGCTGAAAATTCAGTATGGAAT 3541
260 ePheIleMetAsnAsnIleAspGlyGlyGluLeuLeuAsnLeuThrIleGluSerLeuIle 280
3542 TGCGAATTTGGACAAAGATATCTTTTACAGAGCCATCAGAAACAGTGAAGACCTTAA-- 3599
280 AspAspLeuIleIleGluSerLeuGlyLeu-----ArgSerIle 293
3600 -GACTCTTGAAAATTTCAAGACGCTTCTTGTGATTTGATGGAACAGTGAAGGATATG 3658
293 eValLeuArgIleIleGluGlnLeuArgThrIleValIleSerLeuSer----- 309
3659 AATATTA-----TTACTGGAATAATAAGAAAAGACCTTGTCTGTCCAGCGGTACAGTA 3712
310 -----SerGlyIleProAspGluPheIleCysProIleThrArgIle 323
3713 CTTTCTTGACATTTCTCAGATGCTACCA---AGTTTCATCTACTGCTGTGACACAG 3769
323 uLeuMetCysAspProValIleAlaSer 332
3770 ACTGCAAAAGATCTGAGTTTGAATCTCC 3797
Db 3770 ACTGCAAAAGATCTGAGTTTGAATCTCC 3797

RESULT 13
US-09-620-312D-145

/ Sequence 145, Application US/09620312D
/ Patent No. 6569662
/ GENERAL INFORMATION:
/ APPLICANT: Tang, Y. Tom
/ APPLICANT: Liu, Chenghua
/ APPLICANT: Asundi, Vinod
/ APPLICANT: Zhang, Jie
/ APPLICANT: Ren, Feiyang
/ APPLICANT: Chen, Rui-hong
/ APPLICANT: Zhao, Qing A.
/ APPLICANT: Wehrman, Tom
/ APPLICANT: Xue, Aidong J.
/ APPLICANT: Tang, Yonghong
/ APPLICANT: Wang, Jian-Rui
/ APPLICANT: Zhou, Ping
/ APPLICANT: Ma, Yundong
/ APPLICANT: Wang, Dunhui
/ APPLICANT: Wang, Zhiluei
/ APPLICANT: John Tillinghaet
/ APPLICANT: Drmanac, Radoje T.
/ TITLE OF INVENTION: No. 6569662el Nucleic Acids and
/ FILE REFERENCE: 784C1P2B
/ CURRENT APPLICATION NUMBER: US/09/620,312D
/ CURRENT FILING DATE: 2000-07-19
/ PRIOR APPLICATION NUMBER: 09/552,317
/ PRIOR FILING DATE: 2000-04-25
/ PRIOR APPLICATION NUMBER: 09/488,725
/ PRIOR FILING DATE: 2000-01-21
/ NUMBER OF SEQ ID NOS: 1105
/ SOFTWARE: pc_fl_genes Version 1.0
/ SEQ ID NO 145
/ LENGTH: 2130
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (184)..(1236)
US-09-620-312D-145
Alignment Scores:
Pred. No.: 9,6e-17 Length: 2130
Score: 226.50 Matches: 96
Percent Similarity: 41.05% Conservative: 60
Best Local Similarity: 25.26% Mismatches: 137
Query Match: 11.06% Indels: 88
Gaps: 15
US-10-077-111-13 (1-384) x US-09-620-312D-145 (1-2130)
QY 25 LeuAlaThrCysSerLeuAspIleAspGlyLeuTyrSerLeuArgAspPheThrGlu 44
Db 149 CTAGCTACAGCTTTCTGAAGACAAATCATTAATAGTATGAGCATG-----TATGCCGCG 202
45 LeuProHisSerProLeuLeuAspPheThrIleTyrAlaValHisCysCysCysPheSerPro 64
203 CGCTTCCTGATATCTTGTGATGACATACATACACACTGTGACGCTGTCGCA--ATTGATACC 261
65 SerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeuTriPheThrGlu 84
262 GATGGAAGACTAATGTGTGATGATGAGATTAATAATTAATTTGGGATACACACA 321
85 AsnGlyGlnMetLeuAlaValMetGluInProSerGlySerProValArgValCysGln 104
322 AATAGCAATGTGTTAATATATCTTCAGATTCCTGTTGGATTT--GCAATTTTGTGAC 378
105 PheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTyr 124
Db 379 TTTAACCTCAGAGGTATCATGATGATCTTCAGACGCTTCGATCAAACTGTGAAGTCTGG 438
125 AsnAlaGlnSerTyrIleLeuTyrArgCysGlySerValIleAspGlySerLeuAlaIle 144
Db 439 GATGTAAAGTGAACAATAATTAATTAATCAAGCATTAACAGATTCACAGCGGTGAGATTATTC 498


```
Db 716 GTCGCCCATTTTACCTGTGTCGCCCATCTTCAGAAATAATGATGATCTTCAGATTGCG 775
QY 332 SerAspGlyTyrSerTyrGluValMetGluAsnTrpIleSerIlySlySarg 351
   |||||
   |||||
   |||||
Db 776 GCGGATGGTTTCACTTATGAAAGGAGATGCGGGAATGGTTGGACAACGACGACAT 835
QY 352 ThrSerProMetThrThrLeuValLeuProSerAlaValLeuThrProAsnAlaGlnTrpLeu 371
   |||||
   |||||
   |||||
Db 836 ACATCCCAATGACCACTTGAAAGCTTATGATTCATTCATCTCACTCCCAACGCGGTTA 895
QY 372 LysMetAlaIleAsnArgTrpLeu 379
   |||||
   |||||
   |||||
Db 896 CGGTCCGCAATTCAGAGTGGCTC 919

RESULT 15
US-09-690-364-10
; Sequence 10, Application US/09690364
; Patent No. 6468795
; GENERAL INFORMATION:
; APPLICANT: Hong Zhang
; APPLICANT: Andrew T. Walt
; TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
; FILE REFERENCE: RTS-0190
; CURRENT APPLICATION NUMBER: US/09/690,364
; NUMBER OF SEQ ID NOS: 100
; SEQ ID NO 10
; LENGTH: 5152
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (586) ... (4302)
US-09-690-364-10

Alignment Scores:
Pred. No.: 4,83e-15 Length: 5152
Score: 218.00 Matches: 65
Percent Similarity: 39.15% Conservative: 27
Best Local Similarity: 27.66% Mismatches: 75
Query Match: 10.65% Indels: 68
DB: 3 Gaps: 6

US-10-077-111-13 (1-384) x US-09-690-364-10 (1-5152)
QY 3 LysLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPheSerPhe 22
   |||||
   |||||
   |||||
Db 2620 AACCTTGTGACACCTTACGACGACGACCTGGAGCAAGTCAATTGCTGCCATTTACCCAAC 2679
QY 23 Ser-----LeuLeuAlaThrCysSerLeuAspIlyThrIleArgLeuTyrSer 38
   |||||
   |||||
   |||||
Db 2680 AGTAGTAACCACTTCTCTGGCCACCGGGTCAATGATTTCTTCTCAAGCTCTGGGAT 2739
QY 39 LeuArgAspPheThrGluLeuProHisSerProLeuIlyPheHisThrTyrAlaValHis 58
   |||||
   |||||
   |||||
Db 2740 TTGAATCAA-----AAGAAATGCGAAATAACCATGTTGGTGCACACGAACTCAGTCAAC 2793
QY 59 CysCysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrTrp 78
   |||||
   |||||
   |||||
Db 2794 CACTGCAAGTTCTCACCACGACGATGAGCTTGGCTAGCTCAGCTGACGGAACCTTTA 2853
QY 79 ValLeuTrpAsnThrGluAsnGlyGln----- 87
   |||||
   |||||
   |||||
Db 2854 AGGCTTTGGGATGTGATGATCAGCAAAAGAGAGAAAGCATTAATGGAAGCGCTTCTTC 2913
QY 87 ----- 87
Db 2914 CTGAGTTCAAGAGACCTCCAGAGAGATGTGAGGTGATCGTGAAGTGTGTTCTGTGATCT 2973
QY 87 ----- 87
Db 2974 GCAGATGTGCAAAATAATATGTGCGACGAAAAAACAAGTCTCTTTTGTGATTTCAAT 3033
```

```
QY 88 -----MetLeuAlaValMetGluGlnProSerGlySerProValArgValCysGln 104
   |||||
   |||||
   |||||
Db 3034 ACTAGTGGCTTATGTGGCAAGAGATCCACAGGCCATCACAGCAACCATCCAGTACTGTGAC 3093
QY 105 PheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrp 124
   |||||
   |||||
   |||||
Db 3094 TTCTCCCTTAAGACCATTTGGCTGTGATGGCTGTCTGACATGACTGTGTGAGATTGTGG 3153
QY 125 AsnAlaGlnSer---TyrIlyLeuTyrArgCysGlySerValIlyAspGlySerLeuAla 143
   |||||
   |||||
   |||||
Db 3154 AACATAGACTCCCGCCTAAAGGTGGCGGCACTGC-----AGAGACATTGAGT 3201
QY 144 -----AlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
   |||||
   |||||
   |||||
Db 3202 TGGGTTACAGGTGTGTGATGTTTCTCCCGATGGCTCCTCATTTTGTGACAGCTTCTGATGAC 3261
QY 161 GlyAspLeuThrValTrpAspArgIlyMetAspGlyLeuHisSer 175
   |||||
   |||||
   |||||
Db 3262 CAACATAGAGGTCTGGGAGACAAATAAGTATGCAAGACTCT 3306
```

Search completed: February 5, 2005, 14:56:56
Job time : 263 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 08:18:36 ; Search time 4596 Seconds
(without alignments)
10534.759 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gaattcgcttcacctcg.....acaccaaagtaagaattc 1272

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 3966416

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_esc1.*
2: gb_esc2.*
3: gb_hnc.*
4: gb_esc3.*
5: gb_esc4.*
6: gb_esc5.*
7: gb_esc6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length DB | ID | Description |
|------------|-------|-------------|-----------|----|--------------------|
| C 1 | 199 | 15.6 | 199 | 2 | BE669639 7e16e07.x |
| C 2 | 199 | 15.6 | 199 | 2 | BR110272 7n51g07.x |
| C 3 | 156.8 | 12.3 | 160 | 1 | AA501581 ne97g11.s |
| C 4 | 130.2 | 10.2 | 199 | 1 | AA141575 mc35f10.x |
| C 5 | 111.8 | 8.8 | 190 | 2 | BF653223 276923.MA |
| C 6 | 80.8 | 6.4 | 138 | 6 | CA587431 LBRI2034 |
| C 7 | 78.8 | 6.2 | 82 | 1 | AA912881 c127b06.s |
| C 8 | 71.4 | 5.6 | 191 | 2 | BR323974 BR323974 |
| C 9 | 49.2 | 3.9 | 177 | 2 | BR186455 BR186455 |
| C 10 | 45.4 | 3.6 | 102 | 7 | CR455364 CR455364 |
| C 11 | 42.8 | 3.4 | 189 | 2 | AM032239 EST275693 |
| C 12 | 39.8 | 3.1 | 184 | 9 | AG211342 Oryza sat |
| C 13 | 39.2 | 3.1 | 144 | 1 | AT622018 486031E09 |
| C 14 | 34.2 | 2.7 | 175 | 2 | BE240075 BS7404124 |
| C 15 | 33.8 | 2.7 | 195 | 7 | CF370509 IG50B08.Y |
| C 16 | 33.6 | 2.6 | 172 | 1 | AF071906 AF071906 |
| C 17 | 33.6 | 2.6 | 172 | 1 | AF071907 AF071907 |
| C 18 | 33.4 | 2.6 | 165 | 7 | CP968712 CP968712 |
| C 19 | 33.2 | 2.6 | 168 | 9 | CL790258 OR_BB010 |
| C 20 | 33.2 | 2.6 | 168 | 9 | CL790258 OR_BB010 |
| C 21 | 33.2 | 2.6 | 179 | 1 | AV174594 AV174594 |
| C 22 | 33 | 2.6 | 181 | 7 | N89830 ZB34G04.81 |
| C 23 | 33 | 2.6 | 183 | 6 | C01352 HMG5000833 |
| C 24 | 32.6 | 2.6 | 130 | 8 | BZ394951 E1NAJ35TR |

| | | | | | |
|------|------|-----|-----|---|--------------------|
| C 25 | 32.4 | 2.5 | 121 | 2 | AW039819 EST282292 |
| C 26 | 32.4 | 2.5 | 149 | 1 | AF109011 AF109011 |
| C 27 | 32.2 | 2.5 | 171 | 1 | AV289619 AV289619 |
| C 28 | 31.8 | 2.5 | 170 | 8 | AZ721313 PCPI-24-1 |
| C 29 | 31.8 | 2.5 | 171 | 2 | AW022465 dE39b11.y |
| C 30 | 31.8 | 2.5 | 182 | 1 | AA463921 aa08h11.b |
| C 31 | 31.6 | 2.5 | 158 | 1 | AA577442 nm91a11.8 |
| C 32 | 31.6 | 2.5 | 170 | 7 | CF526046 lu-c-c-76 |
| C 33 | 31.2 | 2.5 | 130 | 5 | BP519881 BP519881 |
| C 34 | 31.2 | 2.5 | 148 | 7 | W49852 ZEST00026.M |
| C 35 | 31.2 | 2.5 | 152 | 2 | BF361608 POR-L-103 |
| C 36 | 31.2 | 2.5 | 156 | 6 | CD989424 CD989424 |
| C 37 | 31.2 | 2.5 | 169 | 7 | CV296852 EST185229 |
| C 38 | 31.2 | 2.5 | 189 | 9 | CR044488 CR044488 |
| C 39 | 31 | 2.4 | 151 | 5 | BP699577 BP699577 |
| C 40 | 31 | 2.4 | 158 | 5 | BW154677 BW154677 |
| C 41 | 31 | 2.4 | 178 | 7 | CV427795 RC6-FT013 |
| C 42 | 30.8 | 2.4 | 200 | 1 | AU089064 AU089064 |
| C 43 | 30.8 | 2.4 | 171 | 4 | BG266397 BG266397 |
| C 44 | 30.8 | 2.4 | 186 | 4 | BG272274 BG272274 |
| C 45 | 30.6 | 2.4 | 146 | 7 | CF249442 CF249442 |

ALIGNMENTS

RESULT 1
LOCUS BE669639/c 199 bp mRNA EST 08-SEP-2000
DEFINITION 7e16e07.x1 NCI CGAP Lu24 Homo sapiens CDNA clone IMAGE:328260 3' similar to TR:080828 080828 HYPOTHETICAL 88.8 KD PROTEIN. ;, mRNA sequence.

ACCESSION BE669639
VERSION BE669639.1 GI:10030180
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 199)
AUTHORS NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lemon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL, send email to: info@image.lnl.gov

Trace considered overall poor quality
Seg primer: -40up from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 199

FEATURES

source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_image="328260"
/tissue_type="carcinoid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Lu24"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Plasmid DNA from the normalized library NCI CGAP Lns was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonoids

ORIGIN 1414920-1417991 and 1520904-1522439). Subtraction by Bento Soares and M. Fatima Bonaído. "

Query Match 15.6%; Score 199; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.5e-45;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1056 TTTATATGTCCTAATACTAGAGAACTTATGAAAGATCCGTCATCGCATGAGATGCTAT 1115
DB 199 TTTATATGTCCTAATACTAGAGAACTTATGAAAGATCCGTCATCGCATGAGATGCTAT 140
QY 1116 TCATATGAAAAGAGAGCATGAAAATTTGATCAGCAAAAAGAAAGCTACAGTCCCATG 1175
DB 139 TCATATGAAAAGAGAGCATGAAAATTTGATCAGCAAAAAGAAAGCTACAGTCCCATG 80
QY 1176 ACAAACTTGTCTTCTTCCTGAGCGGTACTTACCAAAATGAGACTCGAAAATGGCCATC 1235
DB 79 ACAAACTTGTCTTCTTCCTGAGCGGTACTTACCAAAATGAGACTCGAAAATGGCCATC 20
QY 1236 AATAGATGCTGAGACAC 1254
DB 19 AATAGATGCTGAGACAC 1

RESULT 2
BF110272/c 199 bp mRNA linear EST 20-OCT-2000
LOCUS 7n51907.x1 NCI CGAP Lu24 Homo sapiens CDNA clone IMAGE:3568309.3,
DEFINITION similar to TR:Q9SYC6 Q9SYC6 HYPOTHETICAL 48.9 KD PROTEIN.;, mRNA
sequence.

ACCESSION BF110272 GI:10939962
VERSION BF110272
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 199)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov

Trace considered overall poor quality
Seq primer: -40UP from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 199
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3568309"
/issue_type="carcinoid"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_Lu24"
/note="Organ: lung; Vector: pTT3D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneids
1414920-1417991 and 1520904-1522439). Subtraction by Bento

ORIGIN Soares and M. Fatima Bonaído. "

Query Match 15.6%; Score 199; DB 2; Length 199;
Best Local Similarity 100.0%; Pred. No. 1.5e-45;
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1044 ATTCTGATGAATTTATATATGTCCTAATGAGAACTTATGAAAGATCCGTCATCGCA 1103
DB 199 ATTCTGATGAATTTATATGTCCTAATGAGAACTTATGAAAGATCCGTCATCGCA 140
QY 1104 TCAGATGCGCTATTCATATGAAAAGAGCAATGCAAAAATTTGATCGCAAAAAGAAAGCT 1163
DB 139 TCAGATGCGCTATTCATATGAAAAGAGCAATGCAAAAATTTGATCGCAAAAAGAAAGCT 80
QY 1164 ACAATGCCATGACAAATCTGTCTTCTTCCTGAGCGGTACTTACCAAAATGAGACTCTG 1223
DB 79 ACAATGCCATGACAAATCTGTCTTCTTCCTGAGCGGTACTTACCAAAATGAGACTCTG 20
QY 1224 AAAATGCCATCATATGAT 1242
DB 19 AAAATGCCATCATATGAT 1

RESULT 3
AA501581/c 160 bp mRNA linear EST 19-AUG-1997
LOCUS ne97g11.s1 NCI CGAP Kid1 Homo sapiens CDNA clone IMAGE:912260
DEFINITION similar to SW:PKM THECU P49695 PUTATIVE SERINE/THREONINE-PROTEIN
KINASE PKM.;, mRNA sequence.

ACCESSION AA501581 GI:2236548
VERSION AA501581
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens (human)

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 160)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
COMMENT Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html

Trace considered overall poor quality
Insert Length: 519 Std Error: 0.00
Seq primer: -40ml3 fwd. BT from Amersham
High quality sequence stop: 1.
Location/Qualifiers
1. 160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:912260"
/issue_type="kidney"
/lab_host="DH10B"
/clone_1lb="NCI_CGAP_Kid1"
/note="Vector: pAMP10; mRNA made from invasive kidney
tumor, cDNA made by oligo-dt priming. Non-directionally
cloned. Size-selected on agarose gel, average insert
size 600 bp. Reference: Krizman et al. (1996) Cancer
Research 56:5380-5383."

ORIGIN

Query Match 12.3%; Score 156.8; DB 1; Length 160;
 Best Local Similarity 98.8%; Pred. No. 1.8e-33;
 Matches 158; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 249 CATTCTCATTGAAGTTTCATCTATGCTGTCCACGTGCTGCTTTCTCCCTTCAGCA 308
 DB 160 CATTCTCATTGAAGTTTCATCTATGCTGTCCACGTGCTGCTTTTATCCCTTCAGCA 101
 QY 309 CATTATTTGGCATGCTGTTCAACAGATGCTACCACTGCTCTATGGAATGAAATGGA 368
 DB 100 CATTATTTGGCATGCTGTTCAACAGATGCTACCACTGCTCTATGGAATGAAATGGA 41
 QY 369 CAGATGCTGGCAGTATGGAACACCTAGTGGCAGCCCTG 408
 DB 40 CAGATGCTGGCAGTATGGAACACCTAGTGGCAGCCCTG 1

RESULT 4
 A1415375 199 bp mRNA linear EST 09-FEB-1999
 LOCUS mc35ff10.x1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
 DEFINITION IMAGE:350539.3 similar to TR:023121 023121 F10G19.3 PROTEIN. ;
 mRNA sequence.
 A1415375
 VERSION A1415375.1 GI:4258879
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 199)
 Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
 Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y.,
 Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R.,
 Ritten, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
 Waterson, R. and Wilson, R.
 The WashU-NCI Mouse EST Project 1999
 Unpublished (1999)
 Contact: Marra M/WashU-NCI Mouse EST Project 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watsn.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@imgc.lnl.gov) for further information.
 This clone was previously sequenced on the 5' end only, this new
 data is from the 3' end
 Trace considered overall poor quality
 Possible reversed clone; similarity on wrong strand
 High quality sequence stop: 1.
 Location/Qualifiers
 1. 199
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:350539"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 /clone_1lb="Soares mouse p3NMF19.5"
 /note="Vector: pRT73D (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5',
 TGTTCACCATCTGAAGTGGAGCGCGCCGATTTTCTTTTCTTTT 3'],
 double-stranded cDNA was size selected, ligated to Eco RI
 adapters (Pharmacia), digested with Not I and cloned into
 the Not I and Eco RI sites of a modified pRT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M. Fatima Bonaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."

FEATURES
 SOURCE

ORIGIN

Query Match 10.2%; Score 130.2; DB 1; Length 199;
 Best Local Similarity 78.4%; Pred. No. 8.6e-26;
 Matches 156; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1044 ATTCTCGATGAATTTATATGCTCCATTAATGAGAACTTATGAAAGTCCGGTCATGCA 1103
 DB 199 ATCCCTGACAGATTCATGTCACAAATTAACAGAACTCATGAGAGACCCGTCATGCA 140
 QY 1104 TCAGATGCTATTCATATGAGAAAGAGCAATGAGAAATGATGACAAAGAAAGT 1163
 DB 139 TCAGATGCTATTCCTTACGAGAGAGAGCAATGCAAGCTGATCCAGAGAGAGCT 80
 QY 1164 ACAAGTCCATGACAAATCTGTTCTTCTTACGCGGTACTTACCACTAATGAGACTCTG 1223
 DB 79 ACCAGCCCATGTCATTAATTTGGCTCTTCATTCATGCTGTAAGACCCCAAGAGACACTG 20
 QY 1224 AAAATGGCCATCAATAGAT 1242
 DB 19 AAGATGGCCATCAACCAAT 1

RESULT 5
 BF653223/c 190 bp mRNA linear EST 25-APR-2001
 LOCUS BF653223
 DEFINITION 276923 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.
 ACCESSION BF653223
 VERSION BF653223.1 GI:11918292
 KEYWORDS EST.
 SOURCE Bos taurus (cow)
 ORGANISM Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 Bovinae; Bos.
 1 (bases 1 to 190)
 Smith, T.P.L., Grose, W.M., Freking, B.A., Roberts, A.J., Stone, R.T.,
 Casas, E., Wray, J.B., White, J., Cho, J., Fahrnenkrug, S.C.,
 Bennett, G.L., Heaton, M.P., Laegreid, W.W., Rohrer, G.A.,
 Chitko-McKown, C.G., Perlea, G., Holt, I., Karanycheva, S., Liang, F.,
 Quackenbush, J. and Keefe, J.W.
 Sequence evaluation of four pooled-tissue normalized bovine cDNA
 libraries and construction of a gene index for cattle
 Genome Res. 11 (4), 626-630 (2001)
 11282978
 Contact: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smith@email.marc.usda.gov
 Single pass sequencing. Bases called and alt trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 18
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCACGTCACACAG
 Plate: 66 row: H column: 24
 Seq primer: ATTAGGTGACACTATG.
 Location/Qualifiers
 1. 190
 /organism="Bos taurus"
 /mol_type="mRNA"
 /db_xref="taxon:9913"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /clone_1lb="MARC 3BOV"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 library made from pooled tissue from marrow, alveolar
 macrophage, ovary, fetal semitendoneous muscle, and fetal
 longissimus muscle."

FEATURES
 SOURCE

ORIGIN

Query Match 8.8%; Score 111.8; DB 2; Length 190;

Best Local Similarity 87.4%; Pred. No. 1,6e-20;
Matches 146; Conservative 0; Mismatches 17; Indels 4; Gaps 2;

QY 1104 TCAGATGGCTATTCATATGAGAAAGCAATGAGAAATTCAGCAAAAGAAAGCT 1163

DB 181 TTGAGAGGGCTACCTATGTAAGAAAGCAATGAGAAATTCAGTACAGAAAGAAAGCT 122

QY 1164 ACAAGTCCCATGACAAATCTGTTCTTCCTT-CAGCGGTACTTACACCAATAGAGCTCT 1222

DB 121 ACCAGTCCCATGACAAATCTGTTCTTCTTACAGTACACTTACTCAATATAGAGCTCT 62

QY 1223 GAAATATGGCTCATATGATGCT---GGAGACACACCAAAAGTAA 1266

DB 61 GCAAAATGGCCATGATGATGCTAACTGAGACACATCAAAATGAA 15

RESULT 6
LOCUS CA587431 138 bp mRNA linear EST 12-JAN-2004

DEFINITION LBE12P34 cDNA from mouse aorta Mus musculus cDNA, mRNA sequence.

ACCESSION CA587431
VERSION CA587431.1 GI:40792674

KEYWORDS EST.
SOURCE Mus musculus (house mouse)

REFERENCE
AUTHORS Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 138)

TITLE Borang, S., Anderson, T., Theil, A., Odeberg, J. and Lundberg, J.
Vascular gene expression in atherosclerotic plaque prone regions
analysed by representational difference analysis

JOURNAL Unpublished (2002)

COMMENT Contact: Andersson Tove
Department of Biotechnology
KTH

Teknikringen 34, plan 6, 100 44 Stockholm, Sweden
Tel: +46 8 790 71 29

Fax: +46 8 245452
Email: tove@biochem.kth.se
Representations (amplified cDNA) from plaque prone regions
Seq primer: CTA TGA CCA TGA TTA CGC CAA G.

FEATURES
Location/Qualifiers

1..138
/organism="Mus musculus"
/mol_type="mRNA"
/strain="ADP6-/- and LPLR-/- on C57BL/6x 129 background"
/db_xref="taxon:10090"

/dev_stage="8 weeks old"
/clone_1lb="CDNA from mouse aorta"
/note="Organ: aorta; Site 1: DpnII; Site 2: DpnII; CDNA
was prepared from whole aorta divided in atherosclerotic
plaque prone regions (aortic arch and abdominal aorta
proximal part) and less plaque prone regions (descending
thoracic aorta and abdominal aorta distal part). CDNA was
fragmented with DpnII, linker ligated and amplified to
generate starting material for representational
difference analysis (RDA). The two cDNA pools were
subjected to iterative RDA subtraction and amplification
to enrich for gene fragments differentially expressed at
early stages of atherosclerosis."

ORIGIN

Query Match 6.4%; Score 80.8; DB 6; Length 138;
Best Local Similarity 81.0%; Pred. No. 1.1e-11;
Matches 94; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1151 CAAAAAGAAAGTACAGTCCATGACAAATCTGTTCTTCAGCGGTACTTACACC 1210

DB 4 CAGAAGAGAGCGTACGAGCCCAATTTGGCTCTCCCTTCACTGATGACCCC 63

QY 1211 AATTAGACTCTGAAAAATGGCCATCAATAGATGGCTTGAGACACCAAAAGTAA 1266

ORIGIN

Query Match 6.4%; Score 80.8; DB 6; Length 138;
Best Local Similarity 81.0%; Pred. No. 1.1e-11;
Matches 94; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

QY 1151 CAAAAAGAAAGTACAGTCCATGACAAATCTGTTCTTCAGCGGTACTTACACC 1210

DB 4 CAGAAGAGAGCGTACGAGCCCAATTTGGCTCTCCCTTCACTGATGACCCC 63

QY 1211 AATTAGACTCTGAAAAATGGCCATCAATAGATGGCTTGAGACACCAAAAGTAA 1266

DB 64 AAACAGACACTGAGATGCGCCATCAACCGATGCTGAGAGACGACAGAACTGAA 119

RESULT 7
LOCUS AA912881/c 82 bp mRNA linear EST 26-AUG-1998

DEFINITION AA912881
IMAGE:1524659 3' similar to TR:023121 023121 F10G19.3 PROTEIN. /,
mRNA sequence.

ACCESSION AA912881
VERSION AA912881
KEYWORDS EST.
SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 82)

REFERENCE
AUTHORS NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncigap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL Contact: Robert Strausberg, Ph.D.
Email: cgarbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 509 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amerham
High quality sequence stop: 1.

FEATURES
Location/Qualifiers

1..82
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1524659"
/lab_host="DH10B"

/clone_1lb="Soares, NFL T_GBC, S1"
/note="Organ: pooled; Vector: pVT73D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung Nbh19w, testis NHT, and B-cell
NCI CGAP GCBI) were mixed, and as circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Query Match 6.2%; Score 78.8; DB 1; Length 82;
Best Local Similarity 97.6%; Pred. No. 3.6e-11;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1044 ATTCCTGATGATTTATATGTCATTAAGTACAGAACTTATGAAAGATCCGTCATGCGA 1103

DB 82 ATTCCTGATGATTTATATGTCATTAAGTACAGAACTTATGAAAGATCCGTCATGCGA 23

QY 1104 TCAGATGGCTATTCATATGAAA 1125

DB 22 TCAGTGGCTATTCATATGAAA 1

ORIGIN

Query Match 6.2%; Score 78.8; DB 1; Length 82;
Best Local Similarity 97.6%; Pred. No. 3.6e-11;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1044 ATTCCTGATGATTTATATGTCATTAAGTACAGAACTTATGAAAGATCCGTCATGCGA 1103

DB 82 ATTCCTGATGATTTATATGTCATTAAGTACAGAACTTATGAAAGATCCGTCATGCGA 23

QY 1104 TCAGATGGCTATTCATATGAAA 1125

DB 22 TCAGTGGCTATTCATATGAAA 1

ORIGIN

Query Match 6.2%; Score 78.8; DB 1; Length 82;
Best Local Similarity 97.6%; Pred. No. 3.6e-11;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1044 ATTCCTGATGATTTATATGTCATTAAGTACAGAACTTATGAAAGATCCGTCATGCGA 1103

DB 82 ATTCCTGATGATTTATATGTCATTAAGTACAGAACTTATGAAAGATCCGTCATGCGA 23

QY 1104 TCAGATGGCTATTCATATGAAA 1125

DB 22 TCAGTGGCTATTCATATGAAA 1

ORIGIN

Query Match 6.2%; Score 78.8; DB 1; Length 82;
Best Local Similarity 97.6%; Pred. No. 3.6e-11;
Matches 80; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 191)

TITLE
JOURNAL
COMMENT

RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216

Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermoregulation and thermoregulation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.riken.go.jp) for further details.

FEATURES
source

1. 191
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="B430102G11"
/sex="male"
/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/clone_id="RIKEN full-length enriched, 4 days neonate male adipose"
Note: Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5']
GAGGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTT 3', cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 229.0. Second strand cDNA was prepared with the primer adapter of sequence [5'] GAGGAGAGATTCGAGTATTAATATATCCCCCCCCCC 3'. cDNA was cleaved with XhoI and BamHI. Vector: a modified Bluescript KS(+) after bulk excision from Lambda FLX I."

ORIGIN
Query Match

5.6%; Score 71.4; DB 2; Length 191;

Best Local Similarity 74.4%; Pred. No. 6.3e-09; Matches 90; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1151 CAAAAAGAACGACGACAGAGTCCCATGCAATCTGTTCTTCCCTTACGCGTACTTACACC 1210
2 CAAAGAGAGACCTACAGACGCCCATGACAGATTTCCTCTCCCTTACGTTGACTGACCCC 61
Db 1211 AATAGACTCTGAAATGCGCCATCATGATGCTGAGACACACCAAAAGTAAAGAT 1270
62 CAACGACACCTGAAGATGCGCTTCAACGATGCTGAGAGACCACGAGAAAGGAGCGT 121
QY 1271 T 1271
Db 122 T 122

RESULT 9
BI186455
LOCUS

BI186455 RIKEN full-length enriched, adult male spinal cord Mus
musculus cDNA clone A330033K06 3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

EST.
Mus musculus (house mouse)

REFERENCE
AUTHORS

1 (bases 1 to 177)
Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci, P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N., Hirose, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M., Izawa, M., Kadoya, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N., Kiyosabe, M., Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tomimaga, N., Toya, T., Tsunoda, Y., Watanabe, S., Yamamura, T., Yamane, I., Yano, R., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M., and Hayashizaki, Y.
RIKEN Mouse ESTs (Konno, H., et al.)
Unpublished (2000)

TITLE
JOURNAL
COMMENT

Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Teurumi-Ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@sc.riken.jp, URL: http://genome.gsc.riken.jp/
Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki, N., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
Thermoregulation and thermoregulation of thermolabile enzymes by trehalose and its application for the synthesis of full length cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
Itoh, M., Katsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J., Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki, Y., and Hayashizaki, Y.
Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
Please visit our web site (http://genome.riken.go.jp) for further details.

FEATURES
source

1. 177
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A330033K06"

```

/sex="male"
/tissue_type="spinal_cord"
/dev_stage="adult"
/lab_host="pdl10b"
/clone_id="Riken full-length enriched, adult male spinal
cord"
/notes="Site 1: Sali; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer 15'
GAGGAGAGAAGGATCCACAGCTCTTTTCTTTTTTTTTTNN 3', cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. cDNA went through one round of normalization
to Rot = 10.0 and subtraction to Rot = 459.0. Second
strand cDNA was prepared with the primer adapter of
sequence 15' GAGGAGAGATTCCTCAGTTATTTATTTATATCCCCCCCCCCCC
3'. cDNA was cleaved with XhoI and BamHI. Vector: a
modified pluescript KS(+) after bulk excision from Lambda
ELC I."

```

ORIGIN

| | | | | |
|-----------------------|-----------------|------------------|-----------|-------------|
| Query Match | 3.9% | Score 49.2; | DB 2; | Length 177; |
| Best Local Similarity | 73.3%; | Pred. No. 0.014; | | |
| Matches 63; | Conservative 0; | Mismatches 23; | Indels 0; | Gaps 0; |

0y 1181 TCTTGTCTCTCCCTTCAGCCGTACTTACACCAAAATGTGAACTCTGAAAATGGCCATCAATG 1240
 Db 18 TTTCCTTTCCCTTCACTACTCCTTCACCCCAACGAGACACTGAAGATGCCATCAACCG 77

```

QY      1241 ATGGCTGGAGACACCAAGTAA 1266
          |||||
Db      78  ATGGCTGGAGACCAAGTAA 103

```

| RESULT 10 | LOCUS | DEFINITION |
|-----------|-------------------|--|
| CR455364 | CR455364 | 102 bp mRNA linear EST 02-JUL-2004 |
| | CR455364 | Day 14 bovine embryos (bcal) Bos taurus CDNA clone |
| | bcal0015a.h.02 5' | , mRNA sequence. |

| | | |
|-----------|------------------|-------------|
| ACCESSION | CR455364 | GI:49623753 |
| VERSION | CR455364.1 | |
| KEYWORDS | EST. | |
| SOURCE | Bos taurus (cow) | |
| ORGANISM | Bos taurus | |

| REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|--------------------|--|-----------------------|--------------------|--------------------|
| 1 (bases 1 to 102) | Renard, J. P., Sreenan, J. P. and Hue, I | Embryonic ESTs (bcat) | Unpublished (2004) | Contact: Renard JP |

Biologie du Développement et Reproduction
INRA

Domaine de Vilvert 78350 Jouy en Josas, FRANCE
Email: renard@jouy.inra.fr

Funding for EST sequencing was provided by an INRA funding (AIP00163). Funding for the cDNA library construction was provided by an EEC contract (BO14-CT95-0190)..

The library was constructed by I. Hue, BDR, INRA, 78350 JOUY-EN-JOSAS. The double stranded cDNA was un-directionally cloned in a TA cloning vector, *bccl* is a SMART cDNA library. Insert size was between 450 and 3500 bp. Sequencing was done from the 5' end of the clone.

Sequence cleaned of vector, adaptor and repetitions. Contact us at sigenasupport@jouy.inra.fr to obtain the chromatogram of this sequence.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .102 |

1. .102

ORIGIN

```

/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="bcal0015a.h.02"
/tissue_type="early elongating embryos (+/- 15 mm)"
/dev_stage="embryon"
/lab_host="DH"
/clone_lib="Day 14 bovine embryos (bcal)"
/notes="vector: pCR2.1 (Invitrogen); Clone distribution :
AGENCE Resource centre, Francois Plumi,
Francois.Plumi@jouy.inra.fr, INRA, CEA Radiobiologie et
Etude du genome (LRGG), Domaine de Vilvert, 78352,
Jouy-en-Josas cedex, FRANCE, +33 (0) 1.34.65.28.02, +33
(0) 1.34.65.22.73"

```

ORIGIN

| | | | | |
|-----------------------|----------------|----------------|----------|------------|
| Query Match | 3.6% | Score 45.4 | DB 7 | Length 102 |
| Best Local Similarity | 87.3% | Pred. No. 0.15 | | |
| Matches 62 | Conservative 0 | Mismatches 6 | Indels 3 | Gaps 1 |

Oy 1199 GGTACTTACACCAATAGAGCTGTGAAATGGCCATCATAGATGGCTG---GAGACACA 1555
 Db 1 GAACTTACACCAATAGAGCTGTGAAATGGCCATCGATCGATGGCTTAAGTAGACACA 60

| | | | |
|----|------|-----------|------|
| OY | 1256 | CCAAAGTAA | 1266 |
| | | | |
| Db | 61 | TCAAAATAA | 71 |

| RESULT 11 | | | | | |
|------------|--|--------|------|--------|-----------------|
| AM032239 | | | | | |
| LOCUS | AM032239 | 189 bp | mRNA | linear | EST 18-MAY-2001 |
| DEFINITION | EST275693 tomato callus, TAMU Lycopersicon esculentum cDNA clone | | | | |
| | CLEC35P22, mRNA sequence. | | | | |

| | | |
|-----------|-------------------------|------------|
| ACCESSION | AM032239 | |
| VERSION | AM032239.1 | GI:5890995 |
| KEYWORDS | EST. | |
| SOURCE | Lycopersicon esculentum | (tomato) |
| ORGANISM | Lycopersicon esculentum | |

REFERENCE
AUTHORS

Eukariota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon. 1 (bases 1 to 189)

Alcalá, J., Vrebalov, J., White, R., Matern, A. L., Vision, T., Holt, E., Liang, F., Upton, J., Craven, M. B., Bowman, C. L., Ann, S., Romling, C. M., Fraser, C. M., Martin, G. B., Tanksey, S. D. and Giovannoni, J.

TITLE Generation of ESTs from tomato callus tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source

Source

```

/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="72A496"
/db_xref="taxon:4081"
/clone="cLEC35P22"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="Hy1-Blue MRF"
/clone_lib="tomato callus, TAMU"
/note="Vector: pBluescript SK(-); Site_1: EcoRI, Site_2:
XhoI; supplier: Giovannoni laboratory; cLEC - Cotyledons
of seedlings 7-10 days post-germination were excised, cut
at both ends and placed on MS medium with no selection.
Mixed callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST library"

```

ORIGIN

Query Match 3.4%; Score 42.8; DB 2; Length 189;
Best Local Similarity 61.8%; Pred. No. 1;
Matches 68; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 1043 AATTCCTGATGAAATTTATATGTCCTCAATAGAGAACTTATGAAAGATCCGGTCATCC 1102
DB 74 AGTACCCCTTATTTATTTGCCCCATTTCTTAAAGATGAAAGATCCGTGCAGAT 133

QY 1103 ATCAGATGCTATTTCATATGAAAGAGCATGGAATTTGATCAGCA 1152
DB 134 CTCACACTGGAGATGACATATGATCGGAAACATCGAAGAAATGATATTCA 183

RESULT 12
AG211342 184 bp DNA linear GSS 09-AUG-2003
LOCUS Oryza sativa (japonica cultivar-group) DNA, clone:NP4028.0.703.1A,
DEFINITION 3' flanking sequence of Tos17 insertion in rice strain NF4028,
genomic survey sequence.
ACCESSION AG211342 GI:32358532
VERSION AG211342
KEYWORDS GSS.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehretoidae; Oryzaceae; Oryza.

REFERENCE 1 Miyao, A., Tanaka, K., Murata, K., Sawaki, H., Takeda, S., Abe, K.,
AUTHORS Shinzuka, Y., Onosato, K. and Hirochika, H.
TITLE Target Site Specificity of the Tos17 Retrotransposon Shows a
Preference for Insertion within Genes and against Insertion in
Retrotransposon-Rich Regions of the Genome
JOURNAL Plant Cell 15 (8), 1771-1780 (2003)
MEDLINE 22779046
PUBMED 12897251

REFERENCE 2 (bases 1 to 184)
AUTHORS Miyao, A., Sawaki, M. and Hirochika, H.
TITLE Direct Submission
JOURNAL Submitted (16-APR-2002) Akio Miyao, National Institute of
Agrobiological Sciences, Molecular Genetics, 2-1-2, Kannondai,
Tsukuba, Ibaraki 305-8602, Japan (E-mail:miyao@affrc.go.jp,
URL: http://cos.nias.affrc.go.jp/, Tel:81-298-38-7020,
Fax:81-298-38-7020)

FEATURES
Location/Qualifiers
source
1..184
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/strain="NP4028"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/clone="NP4028.0.703.1A"
/clone_1lb="PCR product directly amplified from rice
genomic DNA"
/note="The 3' end of retrotransposon Tos17 was found
immediately upstream of this sequence."

ORIGIN
misc_feature
1 /note="This insertion point is base 115457 in AC020666."

Query Match 3.1%; Score 39.8; DB 9; Length 184;
Best Local Similarity 59.1%; Pred. No. 7.3;
Matches 68; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 1066 CAATTAAGTACGAGCTTATGAAAGATCCGATCGATCGATGCTATTATATGAAA 1125
DB 5 CAATATTGAGAGAGTATGAGGATCTCATTTGACGACGATGGTTCAAGTACGAG 64

QY 1126 AGAAGCAATGGAATTTGATCGACAAAAGAAAGCTACAGTCCATGCAAA 1180
DB 65 AGAGGCGCATCAAGAGATGCTCGGACGCGGCGACACACCTCCCGATGCAAA 119

RESULT 13
A1622018/c 144 bp mRNA linear EST 22-APR-1999
LOCUS A1622018
DEFINITION 466031B09.x4 486 - leaf primordia cDNA library from Hake lab Zea
mays cDNA, mRNA sequence.
ACCESSION A1622018
VERSION A1622018
KEYWORDS A1622018.1 GI:4646943
SOURCE EST.
ORGANISM Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoidae; Andropogoneae; Zea.
1 (bases 1 to 144)
Malbot, V.
REFERENCE 1
AUTHORS Maize ESTs from various cDNA libraries sequenced at Stanford
TITLE University
JOURNAL Unpublished (1999)
COMMENT Contact: Malbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: malbot@stanford.edu
Plate: 486031 row: E column: 09.

FEATURES
Location/Qualifiers
source
1..144
/organism="Zea mays"
/mol_type="mRNA"
/cultivar="B73"
/db_xref="taxon:4577"
/tissue_type="leaf primordia"
/dev_stage="P7-p11 leaf"
/lab_host="E.coli XL1-Blue MFR"
/clone_1lb="486 - leaf primordia cDNA library from Hake
lab"
/note="Organ: shoot; Vector: Lambda zap; Hake lab cDNA
library."

ORIGIN
Query Match 3.1%; Score 39.2; DB 1; Length 144;
Best Local Similarity 58.6%; Pred. No. 9.9;
Matches 68; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1041 GGAATTCCTGATGAAATTTATATGTCCTCAATAGAGAACTTATGAAAGATCCGTCATC 1100
DB 143 GCACCTCGAGCCATTTTCATCTGCCCCATCTTCAGAGAGTATGTCATTCATAGCTC 84

QY 1101 GCATCAGATGGCTTATTCATATGAAAGAGCAATGGAATTTGATCAGCAAAA 1156
DB 83 GCGTCAGATGGCTTACACCTACGACGCGACGCAATGAGCTGCTGACGACGAA 28

RESULT 14
BE240075 175 bp mRNA linear EST 12-JUN-2000
LOCUS BE240075
DEFINITION BE240075 MHRP- Medicago truncatula cDNA clone pMHRP-41H18, mRNA
sequence.
ACCESSION BE240075 GI:9056258
VERSION BE240075.1
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoidae; Trifoliaceae;
Medicago.
1 (bases 1 to 175)
Harrison, M. J., Town, C. D., Bowman, C. L., Craven, M. B., Hansen, T. S.,
Holt, I. E., Cho, J. and Fraser, C. M.
ESTs from phosphate-starved roots of Medicago truncatula

JOURNAL COMMENT

Unpublished (2000)
 Contact: Maria J. Harrison
 Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73401, USA
 Tel: 580-223-5810
 Fax: 580-221-7380
 Email: mjharrison@noble.org
 The Samuel Noble Roberts Foundation: N265942e
 TIGR sequence name: MTHAL45TK
 More information is available at...
<http://chryslie.tamu.edu/medicago>
 Seq primer: Skmod (CTA gaa cta gtc gat cc).
 Location/Qualifiers

FEATURES

source

```
1..175
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="A17"
/db_xref="taxon:3880"
/clone="pMHRP-41H18"
/tissue_type="roots"
/dev_stage="phosphate-starved"
/lab_host="XtLR"
/clone_1lb="MHRP-"
/notes="Vector: pBluescript SK-; Site 1: EcoRI, Site 2: XhoI; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this period, they were fertilized twice weekly with 1/2 Hoagland solutions containing 200m potassium phosphate. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-aseIst helper phage and propagated in XtLR cells."
```

ORIGIN

Query Match 2.7%; Score 34.2; DB 2; Length 175;
 Best Local Similarity 53.3%; Pred. No. 2.9e+02;
 Matches 72; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1046 TCCTGATGAATTTATATGTCGAATTAAGTATGAAAGATCCGTCATCCGATC 1105
 |||||
 DB 41 TCCTTCATATTTATTTATGTCGAATCTTCCAGAAAGTCATCGAGATCACAATGATGTC 100
 |||||

QY 1106 AGATGCTATTCATATATAAGAGCAATGAAATTTGATCAGCAAAAAGAACTAC 1165
 |||||

DB 101 AGATGCTTTTATCTTACCAAGCAGAGGCTTCCGAGATGCTTGAACAGTGCATGACGC 160
 |||||

QY 1166 AAGTCCCATGACAAA 1180
 |||||

DB 161 TTCACCGATGACAAA 175
 |||||

RESULT 15

CF370509

LOCUS

DEFINITION

CF370509 195 bp mRNA linear EST 27-AUG-2003
 IG50508.v1 Meloidogyne hapla female SMART pGEM Meloidogyne hapla
 cDNA 5', mRNA sequence.

ACCESSION

VERSION

CF370509.1 GI:34317755

KEYWORDS

SOURCE

ORGANISM

Meloidogyne hapla
 Meloidogyne hapla
 Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
 Tylenchoidea; Heteroderidae; Meloidogyinae; Meloidogyne.

REFERENCE

AUTHORS

1 (bases 1 to 195)
 McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
 Wylie,T., Dence,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
 Bowers,Y., Gibbons,M., Ritzer,E., Bennett,J., Franklin,C.,
 Tsagarisvilli,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
 Underwood,K., Stepcoe,M., Allen,M., Person,B., Swaller,T.,

TITLE
JOURNAL
COMMENT

Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
 McCann,R., Waterston,R. and Wilson,R.
 The Washington Univ. Nematode EST Project, 1999
 Unpublished (1999)
 Contact: McCarter JP
 The Washington Univ. Nematode EST Project, 1999
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 Cloned unidirectionally. Poly(A)⁺ RNA was concentrated and purified
 using Dynabeads (Dyna) and mRNA eluted for first strand synthesis.
 First strand cDNA was created using MMLV RT (Powerscript, Clontech)
 and primed with oligo(dT) with XhoI site and 5' SMART 'anchor' added
 using chimeric DNA-RNA oligo. 12 PCR cycles were done using first
 strand and primers specific to SMART oligo and 3' end. Double
 stranded cDNA was digested using XhoI/NotI, fractionated on
 Chroma-spin 400 columns (Clontech) and ligated to digested
 pGEM-11zf(+) plasmid. Chemically competent DH10B cells were used as
 host cells. Library materials provided by Dr. David Bird of North
 Carolina State University. Library construction by Jeff Rousch. See
www.nematode.net for additional project information.
 Putative full length read
 The vector to vector length is 196
 Seq primer: Sp6.

FEATURES

source

```
1..195
/organism="Meloidogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/tissue_type="whole organism"
/dev_stage="adult"
/lab_host="DH10B"
/clone_1lb="Meloidogyne hapla female SMART pGEM"
/notes="Vector: plasmid (ampicillin resistant); Site 1: XhoI; Site 2: NotI; Cloned unidirectionally. Poly(A)+ RNA was concentrated and purified using Dynabeads (Dyna) and mRNA eluted for first strand synthesis. First strand cDNA was created using MMLV RT (Powerscript, Clontech) and primed with oligo(dT) with XhoI site and 5' SMART 'anchor' added using chimeric DNA-RNA oligo. 12 PCR cycles were done using first strand and primers specific to SMART oligo and 3' end. Double stranded cDNA was digested using XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent DH10B cells were used as host cells. Library materials provided by Dr. David Bird of North Carolina State University. Library construction by Jeff Rousch. See www.nematode.net for additional project information."
```

ORIGIN

Query Match 2.7%; Score 33.8; DB 7; Length 195;
 Best Local Similarity 56.9%; Pred. No. 3.9e+02;
 Matches 62; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

QY 866 AGATTTAAAGATCTTGTTGTTTCAAGATGAATTAACATTTGATGAGAAAGAACTGTT 925
 |||||

DB 58 AATTAATAAAATTTTATTTATATAGTTTCATTTTAAAGAAATTTTAAAGAAATTTTAAAG 117
 |||||

QY 926 GAATCTTACAAAGAAAGTCTGCGTGAATTTGAAATTTGAATCTCTA 974
 |||||

DB 118 AATTTTCAAAATTAATAATTTTCAAGATTTTGAATTTGACTTTTAA 166
 |||||

Search completed: February 5, 2005, 11:39:03
 Job time : 4600 secs

;; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
;; FILE REFERENCE: 107196.132
;; CURRENT APPLICATION NUMBER: US/09/248, 796A
;; CURRENT FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074, 725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096, 409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 13675
;; LENGTH: 189
;; TYPE: DNA
;; ORGANISM: Candida albicans
US-09-248-796A-13675

Query Match 2.2%; Score 28.4; DB 4; Length 189;
Best Local Similarity 48.7%; Pred. No. 79;
Matches 77; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

OY 931 TTACAAAAGAAAGCTGCTGATGATTTGAATACTCTAGACTGCGTAGTAAG 990
DB 183 TTGAAAAAATGATCTGAGAGGGGAGAAAGGGGATTAATTTGAGTTAAAAAAA 124
OY 991 TGCTGAGAAATTTGAAGACTCAGACCAAGTTAAATCCCTTTCTTCAGAAATCTG 1050
DB 123 TAAAGTGAAAAAGTGAAAAAGGTGAAAGGGTGGGTTTGTGTTTGTAAATCCATTTCTT 64
OY 1051 ATGAAATTTATATGTCATTAATCTAGAGAACTTATGAAA 1088
DB 63 ATCTCTATATATTCATTAATTAATTAATTAATA 26

RESULT 3
US-09-513-999C-31162
;; Sequence 31162, Application US/09513999C
;; Patent No. 6783961
;; GENERAL INFORMATION:
;; APPLICANT: Dumas Milne Edwards, J.B.
;; APPLICANT: Duclert, A.
;; APPLICANT: Giordano, J.Y.
;; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
;; Patent No. 6783961
;; FILE REFERENCE: 59.US2.REG
;; CURRENT APPLICATION NUMBER: US/09/513, 999C
;; CURRENT FILING DATE: 2000-02-24
;; PRIOR APPLICATION NUMBER: US 60/122, 487
;; PRIOR FILING DATE: 1999-02-26
;; NUMBER OF SEQ ID NOS: 36681
;; SOFTWARE: Patent.pm
;; SEQ ID NO 31162
;; LENGTH: 194
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-513-999C-31162

Query Match 2.2%; Score 28.2; DB 4; Length 194;
Best Local Similarity 49.7%; Pred. No. 93;
Matches 72; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

OY 860 TGCACAAGATTTAAAAATCTTGTGTAATTTTCAAGATGAATACCTTGATGAAAAAGA 919
DB 1 TGCATCCCATTTTAAAGAAACCTGTAACTTGAAAGATAGAGACTTAACTGATCAAAACAGA 60
OY 920 ACTGTTGAATCTTCAAAAAGAAAGTCTGGCTGATGATTTGAAAATTTGAATCTTAGAGACT 979
DB 61 ACAGATGGAACAGATTTAAAGAAACAGAGCAAACTTACAAATTAAGAAATCTCAAGAGA 120
OY 980 GCGTAGTAAGTCTGAGAAATTT 1004
DB 121 CGCAATGAAGATTAAGAAATCTT 145

RESULT 4

US-09-248-796A-8503/c
;; Sequence 8503, Application US/09248796A
;; Patent No. 6747137
;; GENERAL INFORMATION:
;; APPLICANT: Keith Weinstock et al
;; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
;; FILE REFERENCE: 107196.132
;; CURRENT APPLICATION NUMBER: US/09/248, 796A
;; CURRENT FILING DATE: 1999-02-12
;; PRIOR APPLICATION NUMBER: US 60/074, 725
;; PRIOR FILING DATE: 1998-02-13
;; PRIOR APPLICATION NUMBER: US 60/096, 409
;; PRIOR FILING DATE: 1998-08-13
;; NUMBER OF SEQ ID NOS: 28208
;; SEQ ID NO 8503
;; LENGTH: 183
;; TYPE: DNA
;; ORGANISM: Candida albicans
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (11)
;; OTHER INFORMATION: Identity of nucleotide sequences at the above locations are unkn
US-09-248-796A-8503

Query Match 2.2%; Score 27.6; DB 4; Length 183;
Best Local Similarity 52.2%; Pred. No. 1,4e+02;
Matches 60; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

OY 515 TAGGTGTAAGAAGGCTCCTTGCGGCGCATGATGATTTCTCTTAATGAAGCTTTTGT 574
DB 117 TATTTCAATGCTGATCTGATTTGCTGCATTAATTAATAGTGGCTTAATGATGTTTTC 58
OY 575 CACTGGCTCCTCATGCTGATTTTACAGATGAGTGAATTAATGAGTGTCTG 629
DB 57 TACATCAAGCCATTGATGCTTACACAGAGGCGCATGGAACAAAGTTTCTG 3

RESULT 5
US-08-875-972-27/c
;; Sequence 27, Application US/08875972
;; Patent No. 598564
;; GENERAL INFORMATION:
;; APPLICANT: Huntington Potter and Jinhue Li
;; TITLE OF INVENTION: ASSAY FOR IDENTIFYING GENES CAUSING
;; TITLE OF INVENTION: CHROMOSOME NON-DISJUNCTION
;; NUMBER OF SEQUENCES: 29
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: HAMILTON, BROOK, SMITH & REYNOLDS, P.C.
;; STREET: Two Millitia Drive
;; CITY: Lexington
;; STATE: Massachusetts
;; COUNTRY: USA
;; ZIP: 02173-4799
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patentin Releasee #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/875, 972
;; FILING DATE: 08-AUG-97
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: US 60/002, 448
;; FILING DATE: 16-AUG-1995
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Granahan Esq., Patricia
;; REGISTRATION NUMBER: 32,227
;; REFERENCE/DOCKET NUMBER: HU95-03PA
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (781) 861-6240
;; TELEFAX: (781) 861-9540

INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 200 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-875-972-27

Query Match 2.2%; Score 27.6; DB 2; Length 200;
Best Local Similarity 49.3%; Pred. No. 1.5e+02;
Matches 72; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1056 TTTATATGTCGAATAGAGCACTTATGAAAGATCCGTCATCGCATCAGATGGCGTAT 1115
DB 181 TTTATGGGCTCAAAATTTCTGTGCAAAATTTTGGTCAAGTTTCCATTAAAGATAC 122
QY 1116 TCATATGAAAAAGAAAGCAATGGAATTTGATCAGCAAAAGAAACGTACAGTCCCATG 1175
DB 121 TGATTTTAAACCTAATACTTAAGCTGACACGCAAGAAAAAGAAACCAAGTGTGCC 62
QY 1176 ACAATCTTGTCTTCTTCCTCAGCGGT 1201
DB 61 ACAAAACATCTCTTGTCTTGAGGT 36

RESULT 6
US-09-328-750A-5
Sequence 5, Application US/09328750A
Patent No. 6410722
GENERAL INFORMATION:
APPLICANT: MCGILL University et al.
TITLE OF INVENTION: HUMAN AND MAMMALIAN DNA
REPLICATION ORIGIN CONSENSUS SEQUENCES
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: SMABEY OGILVY RENAULT
STREET: 1981 McGill College Avenue - Suite 1600
CITY: Montreal
STATE: QC
COUNTRY: Canada
ZIP: H3A 2Y3
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/328,750A
FILING DATE: 09-Jun-1999
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/033,374
FILING DATE: 16-DEC-1996
APPLICATION NUMBER: 60/047,322
FILING DATE: 21-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Cui, France
REGISTRATION NUMBER: 4166
REFERENCE/DOCKET NUMBER: 1770-162PCT FC/1d
TELECOMMUNICATION INFORMATION:
TELEPHONE: 514 845-7126
TELEFAX: 514 288-8389
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 91 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-328-750A-5

Query Match 2.1%; Score 27; DB 3; Length 91;
Best Local Similarity 15.7%; Pred. No. 1.4e+02;
Matches 13; Conservative 49; Mismatches 21; Indels 0; Gaps 0;

QY 871 TAAAGATCTTGTGTAATTTCAAGATGATAATACATTGATGAAAGAACTTTGATC 930
DB 8 KRAMRWKDAVMMGAKRWKWHRASACMDKAAKTWKGWTAARRYWKGRKMWWTW 67
QY 931 TTACAAAAGAAAGCTGCGTAT 953
DB 68 KAMSDATYAKWMDAKKMWKRT 90

RESULT 7
US-08-369-829A-13
Sequence 13, Application US/08369829A
Patent No. 5861377
GENERAL INFORMATION:
APPLICANT: Hans Fritz
APPLICANT: Christian Sommerhoff
APPLICANT: Jutta Heim
TITLE OF INVENTION: No. 5861377el Inhibitor
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: No. 5861377artis Corporation
STREET: 59 Route 10
CITY: East Hanover
STATE: New Jersey
COUNTRY: US
ZIP: 07936
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/369,829A
FILING DATE: 6-JAN-95
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EPO 9481006.0
FILING DATE: 7-JAN-94
ATTORNEY/AGENT INFORMATION:
NAME: No. 5861377ak, Henry P.
REGISTRATION NUMBER: 33200
REFERENCE/DOCKET NUMBER: 4-19942/A/DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 277-5110
TELEFAX: (908) 277-4306
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 90 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..90
OTHER INFORMATION: /function= "synthetic oligo for PCR"
US-08-369-829A-13
Query Match 2.1%; Score 26.8; DB 2; Length 90;
Best Local Similarity 61.4%; Pred. No. 1.6e+02;
Matches 43; Conservative 0; Mismatches 27; Indels 0; Gaps 0;

QY 1172 CATGCAAAATCTTGTCTTCAGCGGTACTTACACCAATPAGACTGAAATGCG 1231
DB 6 CACAACCGTTTGTCTTCTTCAACCGTACTTACACGAATACGACAGTAACTTGT 65
QY 1232 CATCAATGA 1241
DB 66 TACAAACACA 75

NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 9338
; LENGTH: 195
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-9338

Query Match 2.1%; Score 26.4; DB 4; Length 195;
Best Local Similarity 54.0%; Pred. No. 3.7e+02;
Matches 54; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 986 TAAAGTCCTGAGGAAATGAAAGCTCAGACCAAGTTAAATCCCTTCTTCAGAAAT 1045
DB 195 TCAAGTGTGAGCAATTCGCGCAATTAACATCCAGAACCAATTTTCTAGTTAAATC 136

QY 1046 TCCGATGATTTATATGTCCTCAATTAAGTAAGAACTTATG 1085
DB 135 TTCTGTGAATTCATTTGATTTATGTCGAGAGTTTGG 96

RESULT 12
US-09-248-796A-14068
; Sequence 14068, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248.796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 14068
; LENGTH: 198
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-14068

Query Match 2.1%; Score 26.4; DB 4; Length 198;
Best Local Similarity 59.2%; Pred. No. 3.7e+02;
Matches 45; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 900 AATTAACATGATGAGAAAGAACTGTTGAATCTTACAAAGAACTGCTGATGATTTG 959
DB 22 AATTAATATGATGGGAAACGCTAGTGAGATGAACAGAAATTTGATGAAGATGATTTG 81

QY 960 AAAATTAATCTCTAG 975
DB 82 GATGATGAATTTGAAG 97

RESULT 13
US-08-535-057A-11
; Sequence 11, Application US/08535057A
; Patent No. 6140104
; GENERAL INFORMATION:
; APPLICANT: LERECUS, DIDIER
; APPLICANT: AGAISSE, HERVE
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES FOR THE CONTROL OF
; TITLE OF INVENTION: THE EXPRESSION OF DNA SEQUENCES IN A CELL HOST
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/535,057A
FILING DATE: 20-DEC-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR94/00525
FILING DATE: 05-MAY-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 93/05387
FILING DATE: 05-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-103-0X PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..144
OTHER INFORMATION: /note= "CORRESPONDS TO NUCLEOTIDES
US-08-535-057A-11
OTHER INFORMATION: 1413 TO 1556 OF SEQ ID NO:1"

Query Match 2.1%; Score 26.2; DB 3; Length 144;
Best Local Similarity 54.7%; Pred. No. 3.5e+02;
Matches 52; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1047 CCGATGAATTTATATGTCCTCAATTAAGAACTTAAGAAATCCGTCATGCATCA 1106
DB 33 CATTAAAAATATATATTCACCGCTTAATGATTAAGAAATCATTTATCAGTTTG 92

QY 1107 GATGCTATTCATATGAAAGAAAGCAATGAGAAA 1141
DB 93 AAAATTAATGATTAATGATTAAGAAAGGAGAGAGAA 127

RESULT 14
US-09-503-252-11
; Sequence 11, Application US/09503252
; Patent No. 6553366
; GENERAL INFORMATION:
; APPLICANT: LERECUS, DIDIER
; APPLICANT: AGAISSE, HERVE
; TITLE OF INVENTION: NUCLEOTIDE SEQUENCES FOR THE CONTROL OF
; TITLE OF INVENTION: THE EXPRESSION OF DNA SEQUENCES IN A CELL HOST
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
; CITY: ARLINGTON
; STATE: VA
; COUNTRY: USA
; ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/503,252
FILING DATE: 14-Feb-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/535,057
FILING DATE: <Unknown>
APPLICATION NUMBER: FR 93/05387
FILING DATE: 05-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: OBLON, NORMAN F.
REGISTRATION NUMBER: 24,618
REFERENCE/DOCKET NUMBER: 660-103-0X PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 144 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..144
OTHER INFORMATION: /note="CORRESPONDS TO NUCLEOTIDES
1413 TO 1556 OF SEQ ID NO:1"
US-09-503-252-11
SEQUENCE DESCRIPTION: SEQ ID NO: 11:

Query Match 2.1%; Score 26.2; DB 4; Length 144;
Best Local Similarity 54.7%; Pred. No. 3.5e+02;
Matches 52; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 1047 CCTGATGAATTTATGTCATTAAGAGAACTTATGAAGATCCGTCGATCA 1106
DB 33 CATTAAAAACATATATTGCAACCGCTAATGATTTATGAAAAATCATTTATCAGTTTG 92

QY 1107 GATGCTATTCATATGAAAGAGCAATGAAAA 1141
DB 93 AAATATATGATATATGATTAAGAAAGAGAGAA 127

RESULT 15
US-09-134-001C-1737/c
Sequence 1737, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 1737
LENGTH: 186
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1737

Query Match 2.1%; Score 26.2; DB 3; Length 186;
Best Local Similarity 58.2%; Pred. No. 4.2e+02;
Matches 46; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 895 AGATGATAATCATTTGATGAAAGAACTGTTGAATCTTACAAAGAAAGCTGGCTGATG 954
DB 93 AAATGAAGACAAAGATTTGATGAGTAACGAATGTTTAAAGTAGTTGAGATTACT 34

QY 955 ATTGAAAAATGAAATCTCT 973
DB 33 ATAGAAAAAGCTTACTCT 15

Search completed: February 5, 2005, 11:43:11
Job time: 247 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 04:21:09 ; Search time 5661 Seconds

(without alignments)
10887.658 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gaattcggttcacacgtcg.....acacacaaagtaagaattc 1272

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: *
1: gb pa: *
2: gb hbg: *
3: gb in: *
4: gb om: *
5: gb ov: *
6: gb pat: *
7: gb ph: *
8: gb pl: *
9: gb pr: *
10: gb ro: *
11: gb sts: *
12: gb sy: *
13: gb un: *
14: gb vi: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------------|--------------------|
| 1 | 1250.2 | 98.3 | 1564 | 9 BC029520 | BC029520 Homo sapi |
| 2 | 954.2 | 75.8 | 1844 | 6 AR339277 | AR339277 Sequence |
| 3 | 912.6 | 71.7 | 1254 | 6 CQ729621 | CQ729621 Sequence |
| 4 | 893.4 | 70.2 | 1919 | 6 AK129983 | AK129983 Homo sapi |
| 5 | 890.2 | 70.0 | 1811 | 6 AX748164 | AX748164 Sequence |
| 6 | 890.2 | 70.0 | 1811 | 6 AK093494 | AK093494 Homo sapi |
| 7 | 878.8 | 69.1 | 1996 | 6 CQ491312 | CQ491312 Sequence |
| 8 | 878.8 | 69.1 | 1996 | 6 CQ497179 | CQ497179 Sequence |
| 9 | 478.6 | 37.6 | 1291 | 6 AX077675 | AX077675 Sequence |
| 10 | 474.6 | 37.3 | 2130 | 10 BC061948 | BC061948 Mus muscu |
| 11 | 436 | 34.3 | 1547 | 10 BC050792 | BC050792 Mus muscu |
| 12 | 430.2 | 34.3 | 1547 | 5 AJ719461 | AJ719461 Gallus ga |
| 13 | 430 | 33.8 | 147021 | 9 AC006501 | AC006501 Homo sapi |
| 14 | 430 | 33.8 | 177089 | 9 AC008277 | AC008277 Homo sapi |
| 15 | 335 | 26.3 | 446 | 6 CQ473162 | CQ473162 Sequence |
| 16 | 331.6 | 26.1 | 366 | 6 CQ096643 | CQ096643 Sequence |
| 17 | 331.6 | 26.1 | 366 | 6 CQ135438 | CQ135438 Sequence |
| 18 | 331.6 | 26.1 | 366 | 6 CQ218732 | CQ218732 Sequence |
| 19 | 331.6 | 26.1 | 366 | 6 CQ294537 | CQ294537 Sequence |

| | | | | | |
|------|-------|------|--------|-------------|--------------------|
| C 20 | 331.6 | 26.1 | 366 | 6 CQ331208 | CQ331208 Sequence |
| C 21 | 329.8 | 25.9 | 409 | 6 CQ482331 | CQ482331 Sequence |
| C 22 | 328.8 | 25.8 | 441 | 6 CQ503452 | CQ503452 Sequence |
| C 23 | 328.8 | 25.8 | 441 | 6 CQ512287 | CQ512287 Sequence |
| C 24 | 297 | 23.3 | 297 | 6 CQ109842 | CQ109842 Sequence |
| C 25 | 297 | 23.3 | 297 | 6 CQ148540 | CQ148540 Sequence |
| C 26 | 297 | 23.3 | 297 | 6 CQ231810 | CQ231810 Sequence |
| C 27 | 297 | 23.3 | 297 | 6 CQ307124 | CQ307124 Sequence |
| C 28 | 297 | 23.3 | 297 | 6 CQ344013 | CQ344013 Sequence |
| C 29 | 276.8 | 21.8 | 245795 | 2 AC096419 | AC096419 Sequence |
| C 30 | 276.8 | 21.8 | 315453 | 2 AC122070 | AC122070 Rattus no |
| C 31 | 275.4 | 21.7 | 196403 | 2 AC102767 | AC102767 Mus muscu |
| C 32 | 275.4 | 21.7 | 209437 | 2 AC111119 | AC111119 Mus muscu |
| C 33 | 275.4 | 21.7 | 217352 | 10 AL929242 | AL929242 Mouse DNA |
| C 34 | 179.6 | 14.1 | 466 | 6 CQ097537 | CQ097537 Sequence |
| C 35 | 179.6 | 14.1 | 466 | 6 CQ136375 | CQ136375 Sequence |
| C 36 | 179.6 | 14.1 | 466 | 6 CQ219679 | CQ219679 Sequence |
| C 37 | 179.6 | 14.1 | 466 | 6 CQ258151 | CQ258151 Sequence |
| C 38 | 179.6 | 14.1 | 466 | 6 CQ295485 | CQ295485 Sequence |
| C 39 | 179.6 | 14.1 | 466 | 6 CQ332079 | CQ332079 Sequence |
| C 40 | 179.6 | 14.1 | 153866 | 9 AC009307 | AC009307 Homo sapi |
| C 41 | 176.8 | 13.9 | 180 | 6 CQ110706 | CQ110706 Sequence |
| C 42 | 176.8 | 13.9 | 180 | 6 CQ149443 | CQ149443 Sequence |
| C 43 | 176.8 | 13.9 | 180 | 6 CQ232723 | CQ232723 Sequence |
| C 44 | 176.8 | 13.9 | 180 | 6 CQ270754 | CQ270754 Sequence |
| C 45 | 176.8 | 13.9 | 180 | 6 CQ308036 | CQ308036 Sequence |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|------|--------|-----------------|
| RESULT 1 | BC029520 | 1564 bp | mRNA | linear | PRI 20-JUL-2004 |
| LOCUS | BC029520 | | | | |
| DEFINITION | Homo sapiens WD repeat and SAM domain containing 1, mRNA (cDNA clone MGC:33855 IMAGE:5301559), complete cds. | | | | |
| ACCESSION | BC029520 | | | | |
| VERSION | BC029520.1 | GI:20810486 | | | |
| KEYWORDS | MGC. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | 1 (bases 1 to 1564) | | | | |
| AUTHORS | Strausberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schaefer, T.E., Brownstein, M.J., Ustin, T.B., Toshlyuk, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Adamson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahney, J., Heulton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalski, U., Smalinski, D.E., Schermer, A., Schein, J.B., Jones, S.J. and Marra, M.A. | | | | |
| TITLE | Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences | | | | |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) | | | | |
| PUBMED | 12477932 | | | | |
| REFERENCES | 2 (bases 1 to 1564) | | | | |
| AUTHORS | Strausberg, R. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (01-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2550, USA | | | | |

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toshiyuki and Piero Carninci (RIKEN)
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-ehgc.stanford.edu>
Contact: (Dickson, Mark) mcd@pax11.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 48 Row: 0 Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 22749102.

FEATURES

source

1.1564
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:33855 IMAGE:5301559"
/issue_type="Brain, hypothalamus"
/clone_lib="NIH MGC_96"
/lab_host="DH10B"
/note="Vector: pBluescript"
1.1564
/gene="WDSAM1"
/note="synonym: FLJ36175"
/db_xref="LocusID:151525"
146..1300
/gene="WDSAM1"
/product="WDSAM1 protein"
/protein_id="AAH29520.1"
/db_xref="GI:20810487"
/db_xref="LocusID:151525"
/translation="MKLIHTLADHDVVNCAFSLSLATSLDTRIRYSRDTPE
LPHSPKFTYAVHCCFSFGHILASCSVDGTTVLMNTENGMALVMEQPSPVAV
CQSPSPSTCLASGADGVVLMNAQSKLYRCGSVXGDSLAACAFAFNSGFYVGS
GDLVWDKMRCKSHKADLGIITCCDFSPQVSDGQSLQIFRLASCQDDCVKWTI
VSEPTDLARTHEHOLKOFTEDESEEDVSTLCAQDLKDLIFRNNINDEKELNLTJK
ESLADLKTESIGLSKVLRIKELRTKVLSSGIPEDEICCTTEMLPVIASDG
YSYKRAMENWISKKRTSPMTLVLPNAVLTFRNRIUKAIAINMLTHQK"

CDS

gene

ORIGIN

Query Match 98.3%; Score 1250.2; DB 9; Length 1564;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1252; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 12 TCACCTGCGGCGACGTGACCCGACCCCGTGGGCACTTGAAGCGGATCCCGGCG 71
DB 47 TCACCTGCGGCGACGTGACCCGACCCCGTGGGCACTTGAAGCGGATCCCGGCG 106
QY 72 CCCCCGCTCTGAGAGCTGTTTTCTTCAATAAAGACAATGGTGAACCTGATTACACA 131
DB 107 CCCCCGCTCTGAGAGCTGTTTTCTTCAATAAAGACAATGGTGAACCTGATTACACA 166
QY 132 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 191
DB 167 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 226
QY 192 TGTCTCTTGGACAAACAAATTCGCTGTACTGTTACGTGATGATGATGATGATGAT 251
DB 227 TGTCTCTTGGACAAACAAATTCGCTGTACTGTTACGTGATGATGATGATGATGAT 286
QY 252 TCTCCATTGAAGTTTCACTTATGCTGTGATGATGATGATGATGATGATGATGAT 311
DB 287 TCTCCATTGAAGTTTCACTTATGCTGTGATGATGATGATGATGATGATGATGAT 346

QY 312 ATTTTGGCATCGTGTTCACAGATGCTACACTGTCTTATGAAATCTGAAAAATGACAG 371
DB 347 ATTTTGGCATCGTGTTCACAGATGCTACACTGTCTTATGAAATCTGAAAAATGACAG 406
QY 372 ATGCTGGCATGATGAAACAGCTTATGAGAGCCCTGTAGAGGTTTGGCAATTTCCCA 431
DB 407 ATGCTGGCATGATGAAACAGCTTATGAGAGCCCTGTAGAGGTTTGGCAATTTCCCA 466
QY 432 GACTCCAGCTGTTGGCATCAGGAGGAGCTGATGAACTGTGGTTTGTGAAATGACAG 491
DB 467 GACTCCAGCTGTTGGCATCAGGAGGAGCTGATGAACTGTGGTTTGTGAAATGACAG 526
QY 492 TCATCAAAATATATAGATGTGTAGTGTAAAGATGCTCTTGGCGGATGTGCATTT 551
DB 527 TCATCAAAATATATAGATGTGTAGTGTAAAGATGCTCTTGGCGGATGTGCATTT 586
QY 552 TCTCCATTGAAGAGCTTTCTTTGTCACTGCTCTCATGTGTGATTTACAGTGTGGAT 611
DB 587 TCTCCATTGAAGAGCTTTCTTTGTCACTGCTCTCATGTGTGATTTACAGTGTGGAT 646
QY 612 GATPAAATGAGGCTGTGCATGATGTAAGAAAGCATGATCTTGGAAATTAAGCTGCTGCAT 671
DB 647 GATPAAATGAGGCTGTGCATGATGTAAGAAAGCATGATCTTGGAAATTAAGCTGCTGCAT 706
QY 672 TTTTCTTCAAGCAGCTTTCTGATGAGAAACAAGTCTTCAATTTTTCGACTGGCATCA 731
DB 707 TTTTCTTCAAGCAGCTTTCTGATGAGAAACAAGTCTTCAATTTTTCGACTGGCATCA 766
QY 732 TGTGTCTAGAGATTTGCCAAGTCAAAATTTGGATTTTCTTTTACCATATCTTAGCAAG 791
DB 767 TGTGTCTAGAGATTTGCCAAGTCAAAATTTGGATTTTCTTTTACCATATCTTAGCAAG 826
QY 792 CGCACAGAAATCAGCTGTAAGAAATTTACGAAATTTGTCAGAGAGGTCGCTCAACA 851
DB 827 CGCACAGAAATCAGCTGTAAGAAATTTACGAAATTTGTCAGAGAGGTCGCTCAACA 886
QY 852 TGGCTTTGTGACACAGATTTTAAAGATCTTGTGTGATTTTCAAGATGATTAATGAT 911
DB 887 TGGCTTTGTGACACAGATTTTAAAGATCTTGTGTGATTTTCAAGATGATTAATGAT 946
QY 912 GAAAAAGAACTGTGAAATCTTACAAAAGAAAGTCTGGCTGATGATTTGAAAAATGAAATCT 971
DB 947 GAAAAAGAACTGTGAAATCTTACAAAAGAAAGTCTGGCTGATGATTTGAAAAATGAAATCT 1006
QY 972 CTAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1031
DB 1007 CTAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1066
QY 1032 CTTTCTTCAAGAAATTCCTGATGAAATTTATGTGCAATTAAGTAACTAGAACTTATGAAAGAT 1091
DB 1067 CTTTCTTCAAGAAATTCCTGATGAAATTTATGTGCAATTAAGTAACTAGAACTTATGAAAGAT 1126
QY 1092 CCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
DB 1127 CCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1186
QY 1152 AAAAAAGAACTGTAAGTCCCATGACCAATCTTGTCTTCTTCAAGCGGTAAGTAACTTACACA 1211
DB 1187 AAAAAAGAACTGTAAGTCCCATGACCAATCTTGTCTTCTTCAAGCGGTAAGTAACTTACACA 1246
QY 1212 AATAGACTCTGAAAAATGCGCATCAATAGATGGCTGGAGACACACCAAAAGTAA 1266
DB 1247 AATAGACTCTGAAAAATGCGCATCAATAGATGGCTGGAGACACACCAAAAGTAA 1301

RESULT 2
AR339277
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE

AR339277 1844 bp DNA linear PAT 17-AUG-2003
Sequence 768 from patent US 6569662.
AR339277.1 GI:33726134
Unknown.

ORGANISM

Unknown:
Unclassified.

1 (bases 1 to 1844)

AUTHORS Tang, Y.T., Zhou, P. and Drmanac, R.T.

TITLE Nucleic acids and polypeptides

JOURNAL Patent: US 6569662-A 78 27-MAY-2003;

FEATURES

SOURCE 1. .1844
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Query Match 75.8%; Score 964.2; DB 6; Length 1844;

Best Local Similarity 81.8%; Pred. No. 2.6e-250;

Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

12 TCACCTCGGGGACGTCGACCCGACCCCGTGGGCACTTGAAGCGATCCCGGCG 71
 65 TCACCTCGGGGACGTCGACCCGACCCCGTGGGCACTTGAAGCGATCCCGGCG 124
 72 CCCCCGCTCGAGGCGTCTTCTCAATAAAGACATGGAACTGATTGACACA 131
 125 CCCCCGCTCGAGGCGTCTTCTCAATAAAGACATGGAACTGATTGACACA 184
 132 TTAGCTGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 191
 185 TTAGCTGATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 244
 192 TGCTCTTGACAAACAAATTCGCTGATGTCGATGTCGATGTCGATGTCGATGTC 251
 245 TGCTCTTGACAAACAAATTCGCTGATGTCGATGTCGATGTCGATGTCGATGTC 304
 252 TCTCCATGTAAGTTCATACCTATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 311
 305 TCTCCATGTAAGTTCATACCTATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 364
 312 ATTTGGCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 371
 365 ATTTGGCATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 424
 372 ATGCTGCAATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 431
 425 ATGCTGCAATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 484
 432 GACTCCAGTGTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 491
 485 GACTCCAGTGTGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 544
 492 TCATCAAAATTAATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 551
 545 TCATCAAAATTAATGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 604
 552 TCTCCTAATGTAAGTTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 611
 605 TCTCCTAATGTAAGTTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 664
 612 GATCAAAATGAGTTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 671
 665 GATCAAAATGAGTTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 724
 672 TTTTCTCAGCAGGCTTCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 731
 725 TTTTCTCAGCAGGCTTCTGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 784
 732 TGTGTCAGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 784
 785 TGTGTCAGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 844
 785 ----- 784
 845 GAATTAATAATTAATAAGTACGATGTCGATGTCGATGTCGATGTCGATGTCGATGTC 904
 785 ----- 784

PD 2/19/00

2

Db 905 TCCCATGATGGGACAGATGCTAGTCTCAGGCTCAGTGATTAAGTCTGTATATATATGAT 964
 785 ----- 784
 Db 965 ACTAATTAATGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1024
 785 ----- 784
 Db 1025 GCTTTTGACCTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1084
 785 ----- 784
 Db 1085 TGGCAATTTGACCTGGAACACCTTTGCAAGGACGAGGACGACGATGATGATGATGAT 1144
 816 TTTACGGAATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 875
 1145 TTTACGGAATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGATGTCGAT 1204
 876 GATCTGTTGATATTTTCAAGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 935
 1205 GATCTGTTGATATTTTCAAGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 1264
 936 AAGAAAGTCTGCTGATGATTTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 995
 1265 AAGAAAGTCTGCTGATGATTTGAATTAATTAATTAATTAATTAATTAATTAATTAAT 1324
 996 AAGAAATTTGAAGTCTGATGATTTGAATTAATTAATTAATTAATTAATTAATTAAT 1055
 1325 AAGAAATTTGAAGTCTGATGATTTGAATTAATTAATTAATTAATTAATTAATTAAT 1384
 1056 TTTATATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1115
 1385 TTTATATGTCATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1444
 1445 TCATATGAAGAAAGACGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1504
 1176 ACAATTTGTCCTTCTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 1235
 1505 ACAATTTGTCCTTCTGATGATGATTAATTAATTAATTAATTAATTAATTAATTAAT 1564
 1236 AATGATGCTGAGACACACCAAAAGTAA 1266
 1565 AATGATGCTGAGACACACCAAAAGTAA 1595
 Db 1565 AATGATGCTGAGACACACCAAAAGTAA 1595

RESULT 3
 CQ729621 1254 bp DNA linear PAT 03-FEB-2004
 DEFINITION Sequence 15555 from Patent WO02068579.
 ACCESSION CQ729621
 VERSION CQ729621.1 GI:42301208
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1
 REFERENCE Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
 AUTHORS Kils, such as nucleic acid arrays, comprising a majority of
 TITLE humanexons or transcripts, for detecting expression and other uses
 JOURNAL Patent: WO 02068579-A 15555 06-SEP-2002;
 FEATURES PE Corporation (NY) (US)
 SOURCE 1. .1254
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 785 ----- 784

| Query Match | Similarity | 71.7% | Score 912.6 | DB 6 | Length 1254 |
|--------------|--------------|--|--------------------|------------|-------------|
| Beat Local | Similarity | 86.5% | Pred. No. 2.5e-236 | | |
| Matches 1092 | Conservative | 0 | Mismatches 54 | Indels 117 | Gaps 3 |
| QY | 111 | ATGTGAAACGTGATTCACACATTTAGCTGATCATGTGTGACGATGTCACTGCTGTGCCTTC | 170 | | |
| Db | 1 | ATGTGAAACGTGATTCACACATTTAGCTGATCATGTGTGACGATGTCACTGCTGTGCCTTC | 60 | | |
| QY | 171 | TCCTTTTCCCTCTTGGCTTACTTGTCTCTTTGGACAAACAAATTGGCTGTACTCGTTACGT | 230 | | |
| Db | 61 | TCCTTTTCCCTCTTGGCTTACTTGTCTCTTTGGACAAACAAATTGGCTGTACTCGTTACGT | 120 | | |
| QY | 231 | GACCTTTACGAACTGCACTGATTCATTCATTTGAAAGTTTCATCTGATGCTGTCACTGCTGC | 290 | | |
| Db | 121 | GACCTTTACGAACTGCACTGATTCATTCATTTGAAAGTTTCATCTGATGCTGTCACTGCTGC | 180 | | |
| QY | 291 | TGTTTCTCCCTTTCAGGACATATTTTGGCATCTGTTTCAACGATGATGACACTGTCTCTA | 350 | | |
| Db | 181 | TGTTTCTCCCTTTCAGGACATATTTTGGCATCTGTTTCAACGATGATGACACTGTCTCTA | 240 | | |
| QY | 351 | TGGAATCTGAAATGGAACAGATGCTGCACTGATGAAACAGCTTATGGACGCTGTG | 410 | | |
| Db | 241 | TGGAATCTGAAATGGAACAGATGCTGCACTGATGAAACAGCTTATGGACGCTGTG | 300 | | |
| QY | 411 | AGGGTTTCCGAGTTTCCCGACACTCCAGCTGTTGGATCAGGGGGACGTGATGGAAC | 470 | | |
| Db | 301 | AGGGTTTCCGAGTTTCCCGACACTCCAGCTGTTGGATCAGGGGGACGTGATGGAAC | 360 | | |
| QY | 471 | GTGTTTTGTGGAATGCAACAGTCATACAAATTAATAGATGTGATGTTTAAAGATGGC | 530 | | |
| Db | 361 | GTGTTTTGTGGAATGCAACAGTCATACAAATTAATAGATGTGATGTTTAAAGATGGC | 420 | | |
| QY | 531 | TCCTTGGGGCATGTGCATTTTCTCTTATGGAAGCTTCTTTGTCACTGGCTCTCATGT | 590 | | |
| Db | 421 | TCCTTGGGGCATGTGCATTTTCTCTTATGGAAGCTTCTTTGTCACTGGCTCTCATGT | 480 | | |
| QY | 591 | GGTATTTAAACAGTGTGGGATGATTAATATGAGTGTCTGCATATGTAAGAAAGACATAT | 650 | | |
| Db | 481 | GGTATTTAAACAGTGTGGGATGATTAATATGAGTGTCTGCATATGTAAGAAAGACATAT | 540 | | |
| QY | 651 | CTTGAATTTACCTGCTGTGATTTTCTTACAGCAGCTTCTGATGAGAAACA----- | 704 | | |
| Db | 541 | CTTGAATTTACCTGCTGTGATTTTCTTACAGCAGCTTCTGATGAGAAACA----- | 600 | | |
| QY | 705 | ----- | 704 | | |
| Db | 601 | AAAAGTACATGAGTGGGCACTGTGCTCTGTCTGCTGTGCTTTTCCCATGATGG | 660 | | |
| QY | 705 | ----- | 725 | | |
| Db | 661 | CAGATGTAGTCTCAGGGATGTCAACAATTGTGCTTTTGAACCTTAACCTTTTACTT | 720 | | |
| QY | 726 | GCAATCATGTGTCAGGATTTGCAAGTCAAAATTTGG---ATTGTTTCTTTAACCATAATC | 782 | | |
| Db | 721 | GCTACTGTTCATATGACAAACAAACATGAAACATCTGGCAATTTGACCTGGAACACTTTGC | 780 | | |
| QY | 783 | TTAGCAAGGCGCACAGAAATCATGCTGTAAGCAATTTACGAAGATTTGTCAGAGAGGCTC | 842 | | |
| Db | 781 | CAAGCAAGGACACAGAAATCATGCTGTAAGCAATTTACGAAGATTTGTCAGAGAGAGAT | 840 | | |
| QY | 843 | GTCTCAACATGGCTTTGTGCAACAAGATTTAAAGATCTTGTGGTATTTTCAAGATGAT | 902 | | |
| Db | 841 | GTCTCAACATGGCTTTGTGCAACAAGATTTAAAGATCTTGTGGTATTTTCAAGATGAT | 900 | | |
| QY | 903 | AACATTTGTAAGAAAGAACTGTGTAATTTTCAAAAGAAAGCTGTGGCTGATTTGAA | 962 | | |
| Db | 901 | AACATTTGTAAGAAAGAACTGTGTAATTTTCAAAAGAAAGCTGTGGCTGATTTGAA | 960 | | |
| QY | 963 | ATTGAATCTTAGGACTGCTGATGTAAGATGCTGAGAAATTTGAAGACTCAGACCAAG | 1022 | | |
| Db | 961 | ATT-----GGACTGCTGATGTAAGATGCTGAGAAATTTGAAGACTCAGACCAAG | 1011 | | |
| QY | 1023 | GTTAATTCCTTTCTTACGAAATTCCTGATGAAATTTATATGTCCATTAAGTACGAAACTT | 1082 | | |

| | | | |
|----|------|---|-------|
| Db | 1012 | GTTHAAATCCCTTCTTCAGGAATTCCTCGATGAAATTTATGTCCTCAATTAACCTAGGAACCTT | 10711 |
| Oy | 1083 | ATGGAAGAATCCGGTCATCGCATCAGATGGCGTAATTCATATGAAAAGGAAGCAATGGAAAT | 11422 |
| Db | 1072 | ATGAAAGATCCGGTCATCGCATCAGATGGCGTAATTCATATGAAAAGGAAGCAATGGAAAT | 11311 |
| Oy | 1143 | TGATATCGCAAAAAAGAAACGTACAAAGTCCCATGCAAAATCTTGTTCTTCTTCAGCGTA | 12022 |
| Db | 1132 | TGATATCGCAAAAAAGAAACGTACAAAGTCCCATGCAAAATCTTGTTCTTCTTCAGCGTA | 11911 |
| Oy | 1203 | CTTACACCAATATAGGACTCTGTGAAAAATGGCCATCATATGATGGCGTGGAGACACACCAAAAG | 12622 |
| Db | 1192 | CTTACACCAAAATAGGACTCTGTGAAAAATGGCCATCATATGATGGCGTGGAGACACACCAAAAG | 12511 |
| Oy | 1263 | TAA 1265 | |
| Db | 1252 | TAA 1254 | |

| RESULT 4 | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS |
|----------|----------|--|-----------|---------|-----------------|--------|----------|-----------|---------|
| AK129983 | AK129983 | 1919 bp | mRNA | linear | PRI 10-SEP-2003 | | | | |
| | | Homo sapiens cDNA FLJ26473 fis, clone KDN04521. | | | | | | | |
| | | AK129983 | | | | | | | |
| | | AK129983.1 GI:34526702 | | | | | | | |
| | | cDNA capping, 5' (full insert sequence). | | | | | | | |
| | | Homo sapiens (human) | | | | | | | |
| | | Homo sapiens | | | | | | | |
| | | Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | | | |
| | | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | | | | |
| 1 | | Teshiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., | | | | | | | |

TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science,

COMMENT
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction and 5'-end one pass sequencing: Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure, Human Genome Center; 3'-end one pass sequencing: RAB; clone selection for full insert sequencing: RAB and Helix Research Institute.
Location/Qualifiers

```

source
1. .1919
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon.9606"
/clone="KDN04521"
/tissue_type="Kidney"
/clone_lib="KDN"
/note="cloning vector: pME18SFL3"

```

| | | | | |
|-----------------------|--------------|---|--------------|-------------|
| Query Match | 70.2% | Score 893.4 | DB 9 | Length 1919 |
| Best Local Similarity | 80.8% | Pred. No. 4.3e-211 | | |
| Matches 1183 | Conservative | 0 | Mismatches 6 | Indels 276 |
| | | | Gaps | 1 |
| Qy | 78 | CTCCTGAGGCTGTTTTCTTCAAAATPAAGAACATGSGTAAACGATTCAACATTAGCT | 137 | |
| | | | | |
| Db | 225 | CCCTTGGGGCTTTTTCTTCAAAATPAAGAACATGSGTAAACGATTCAACATTAGCT | 284 | |

QY 138 GATCATGATGACGATGTCACCTGCTGACCTCTCTCTTCCCTCTGAGTCTACTGCTCC 197
 |||||
 Db 285 GATCATGATGACGATGTCACCTGCTGACCTCTCTCTTCCCTCTGAGTCTACTGCTCC 344
 |||||
 QY 198 TTGGACAAAACAATTCGCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 257
 |||||
 Db 345 TTGGACAAAACAATTCGCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 404
 |||||
 QY 258 TTGAAGTTTCAATCTGCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 317
 |||||
 Db 405 TTGAAGTTTCAATCTGCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 464
 |||||
 QY 318 GCATCGTTCACAGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 377
 |||||
 Db 465 GCATCGTTCACAGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 524
 |||||
 QY 378 GCATGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 437
 |||||
 Db 525 GCATGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 584
 |||||
 QY 438 ACCTGTTTGGCATGAGGCGCATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 497
 |||||
 Db 585 ACCTGTTTGGCATGAGGCGCATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 644
 |||||
 QY 498 AAATTTATATAGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 557
 |||||
 Db 645 AAATTTATATAGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 704
 |||||
 QY 558 AATGGAAGCTCTTCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 617
 |||||
 Db 705 AATGGAAGCTCTTCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 764
 |||||
 QY 618 ATGAGGTGTCTGCAATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 677
 |||||
 Db 765 ATGAGGTGTCTGCAATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 824
 |||||
 QY 678 TCACAGCAGTCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 737
 |||||
 Db 825 TCACAGCAGTCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 884
 |||||
 QY 738 CAGATGTCAGTCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 784
 |||||
 Db 885 CAGATGTCAGTCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 944
 |||||
 QY 785 ----- 784
 |||||
 Db 945 AAATATATAAGTCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1004
 |||||
 QY 785 ----- 784
 |||||
 Db 1005 GATGCGCAGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1064
 |||||
 QY 785 ----- 784
 |||||
 Db 1065 ACTGAGATATATCTTCAACATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1124
 |||||
 QY 785 ----- 784
 |||||
 Db 1125 GCACCTAATACCTTTTCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1184
 |||||
 QY 785 ----- 821
 |||||
 Db 1185 TTGAAGTTTCAATCTGCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1244
 |||||
 QY 822 GAAGATGTCAGAGAGTGTCTCAACATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 881
 |||||
 Db 1245 GAAGATGTCAGAGAGTGTCTCAACATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1304
 |||||
 QY 882 GTTGTATTTTCAAGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 941
 |||||
 Db 1305 GTTGTATTTTCAAGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1364
 |||||
 QY 942 AGTGTGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1001
 |||||

Db 1365 AGTGTGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1424
 |||||
 QY 1002 ATGGAAGCTCTGAGACCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1061
 |||||
 Db 1425 ATGGAAGCTCTGAGACCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1484
 |||||
 QY 1062 TGTCAATTAATCTGAGACCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1121
 |||||
 Db 1485 TGTCAATTAATCTGAGACCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1544
 |||||
 QY 1122 GAAAGGACCAATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1181
 |||||
 Db 1545 GAAAGGACCAATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1604
 |||||
 QY 1182 CTGTTCTTCTGAGACCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1241
 |||||
 Db 1605 CTGTTCTTCTGAGACCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1664
 |||||
 QY 1242 TGGCTGAGACCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1266
 |||||
 Db 1665 TGGCTGAGACCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 1729
 |||||

RESULT 5
 AX748164 1811 bp mRNA linear PAT 20-JUN-2003
 LOCUS AX748164
 DEFINITION Sequence 1689 from Patent EP1308459.
 ACCESSION AX748164
 VERSION AX748164.1 GI:32132552
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1
 Iosgai, T., Sugiyama, T., Otsuki, T., Wakematsu, A., Sato, H., Ishii, S.,
 Yamamoto, J. I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R.,
 Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahashi, K. and
 Masuno, Y.
 Full-length cDNA sequences
 Patent: EP 1308459-A 1689 07-MAY-2003;
 Helix Research Institute (JP) ; Research Association for
 Biotechnology (JP)
 FEATURES
 Location/Qualifiers
 source
 1..1811
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 ORIGIN
 Query Match 70.0%; Score 890.2; DB 6; Length 1811;
 Best Local Similarity 80.9%; Pred. No. 3.2e-220;
 Matches 1178; Conservative 0; Mismatches 3; Indels 276; Gaps 1;
 QY 86 GCGTGTCTTCTCAATAAAGAACATGGAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 145
 |||||
 Db 125 GCGTGTCTTCTCAATAAAGAACATGGAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 184
 |||||
 QY 146 TGAAGATGCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 205
 |||||
 Db 185 TGAAGATGCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 244
 |||||
 QY 206 AACATATGCGCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 265
 |||||
 Db 245 AACATATGCGCTGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 304
 |||||
 QY 266 TCAATCTATGCTGCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 325
 |||||
 Db 305 TCAATCTATGCTGCAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 364
 |||||
 QY 326 TTCAACAGATGTAAGTGTGACCTTTTCTGAACTGCAATTCCTCA 385
 |||||

Db 365 TTCACAGATGATGACCACTGTCTTATGGAATACTGAAATGACAGATGCTGGCAGTAT 424
QY 386 GGAACAGCCTAGTGGAGCCCTGTGAGGGTTTGGCAGATTTCCTCCAGACTCCAGCTGTT 445
Db 425 GGAACAGCCTAGTGGAGCCCTGTGAGGGTTTGGCAGATTTCCTCCAGACTCCAGCTGTT 484
QY 446 GGCATCAGGGGAGCTGATGGAATCTGTGTTTGTGGAAATGACAGTCAATCAAAATTAT 505
Db 485 GGCATCAGGGGAGCTGATGGAATCTGTGTTTGTGGAAATGACAGTCAATCAAAATTAT 544
QY 506 TATATGTGTAAGTGTAAAGATGGCTCTTGGGGGACATGTCATTTTCTCTTAATGGAAG 565
Db 545 TATATGTGTAAGTGTAAAGATGGCTCTTGGGGGACATGTCATTTTCTCTTAATGGAAG 604
QY 566 CTCTTTGTATCAGTGGCTCTCATGTGGTATTTAACAGTGTGGAGATGAATAAATGAGTG 625
Db 605 CTCTTTGTATCAGTGGCTCTCATGTGGTATTTAACAGTGTGGAGATGAATAAATGAGTG 664
QY 626 TCTGCATAGTGAATAAGACATGATCTTGGAAATTAACCTGTGGATTTTCTTCCAGGCC 685
Db 665 TCTGCATAGTGAATAAGACATGATCTTGGAAATTAACCTGTGGATTTTCTTCCAGGCC 724
QY 686 AGTTTCTGATGGAACCAAGTCTTCACTTTTTCAGTCTGGCATCATGTGTCAGATTTG 745
Db 725 AGTTTCTGATGGAACCAAGTCTTCACTTTTTCAGTCTGGCATCATGTGTCAGATTTG 784
QY 746 CCAAGTCAAAATTTGGATTGTTCTTTTACCATATCTT----- 784
Db 785 CCAAGTCAAAATTTGGATTGTTCTTTTACCATATCTTATGTTGAATTAATAATATA 844
QY 785 ----- 784
Db 845 AAGTACACTGATGAGGGGACTGCTCTCTGTTCTGCTTGTGCTTTTCCCATGATGGGCA 904
QY 785 ----- 784
Db 905 GATGCTAGTCTCAGGCTCAGTGATGATGCTGTCTATGATATGATATTAATCTAGTGA 964
QY 785 ----- 784
Db 965 TATACTTCACACTGATCTCAGACACACAGGTATGTCACAACTTGTGCTTTTGACCTTA 1024
QY 785 ----- 784
Db 1025 TACCTTTTACTGCTACTGTTCAATGACAAACAGTGAACATCTGGCAATTTGACT 1084
QY 785 -----AGCAAGGGGACACAGAACATGACGTGAACCAATTTACCGAAGATTG 829
Db 1085 GGAACACTTTTGGCAAGGACAGACAGAACATCAGCTGAACCAATTTTACCGAAGATTG 1144
QY 830 GTCCAGAGAGTGTCTCAACATGAGCTTTGTGCAAAAGATTTTAAAGATCTTGTGGTAT 889
Db 1145 GTCCAGAGAGAGTGTCTCAACATGAGCTTTGTGCAAAAGATTTTAAAGATCTTGTGGTAT 1204
QY 890 TTTCAAGATGATTAACATTTGATGAGAAAAGAACTGTTGAATCTTTACAAAGAAAGTCTGC 949
Db 1205 TTTCAAGATGATTAACATTTGATGAGAAAAGAACTGTTGAATCTTTACAAAGAAAGTCTGC 1264
QY 950 TGTATGATTTGAAAATTTGAATCTCTAGAGACTGCGTAGTAAGTCTGAGGAAAATTTGAAGA 1009
Db 1265 TGTATGATTTGAAAATTTGAATCTCTAGAGACTGCGTAGTAAGTCTGAGGAAAATTTGAAGA 1324
QY 1010 GCTCAGAGACCAAGTAAATCCCTTTCTTCAAGGAATCTGATGAATTTATATGTCGAAT 1069
Db 1325 GCTCAGAGACCAAGTAAATCCCTTTCTTCAAGGAATCTGATGAATTTATATGTCGAAT 1384
QY 1070 AACTAGAGAACTTATGAGAAAGTCCGCTCATGCAATCAGATGCGTATTCATATGAGAAAGA 1129
Db 1385 AACTAGAGAACTTATGAGAAAGTCCGCTCATGCAATCAGATGCGTATTCATATGAGAAAGA 1444
QY 1130 AGCAATGAGAAAATTTGATCAGCAAAAAGAAAGTCAAGTCCCATGCAAAATCTTGTCT 1189
Db 1445 AGCAATGAGAAAATTTGATCAGCAAAAAGAAAGTCAAGTCCCATGCAAAATCTTGTCT 1504

QY 1190 TCCTTCAGCGGTACTTACACCAAAATAGACTCTGAAAATGGCCATCAATAGATGGCTGGA 1249
Db 1505 TCCTTCAGCGGTACTTACACCAAAATAGACTCTGAAAATGGCCATCAATAGATGGCTGGA 1564
QY 1250 GACACACCAAAAGTAAA 1266
Db 1565 GACACACCAAAAGTAAA 1581
RESULT 6
AK093494
LOCUS
DEFINITION
Homo sapiens cDNA FLJ36175 f18, clone TEST12026453, weakly similar
to PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKNA (EC 2.7.1.-).
ACCESSION
AK093494
VERSION
AK093494.1 GI:21752384
KEYWORDS
Oligo capping; f18 (full insert sequence).
SOURCE
Homo sapiens
Organism
Homo sapiens
REFERENCE
AUTHORS
Oca, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R.,
Wakamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H.,
Sekine, M., Obauchi, M., Nishi, T., Shibahara, T., Tanaka, T.,
Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y.,
Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M.,
Shiratori, A., Sudo, H., Hosotani, T., Kaku, Y., Kodaira, H., Kondo, H.,
Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T.,
Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K.,
Tanikawa, M., Yamazaki, M., Minomura, K., Ishida, T., Yamashita, H.,
Muraoka, K., Fujimori, K., Tanai, H., Kimura, M., Watanabe, M.,
Hirooka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S.,
Yosida, M., Hotta, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A.,
Hara, H., Tanabe, T., Nomura, Y., Togiyasu, S., Komai, F., Hara, R.,
Takeuchi, K., Arita, M., Imose, N., Mitsuhashi, K., Yuuki, H., Oshima, A.,
Sasaki, N., Aotsuka, S., Yoshikawa, Y., Matsumura, H., Ichihara, T.,
Shiohara, N., Sano, S., Moriya, S., Momiyama, H., Satoh, N., Takami, S.,
Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H.,
Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T.,
Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K.,
Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M.,
Tashiro, H., Tanigami, A., Fujiwara, T., Ono, T., Yamada, K., Fujii, Y.,
Ozaki, K., Hiro, M., Ohmori, Y., Kawabata, A., Hikiji, T., Kobatake, N.,
Inagaki, H., Ikema, Y., Okamoto, S., Ohtani, R., Kawakami, T.,
Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsumura, K.,
Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togaishi, T.,
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J.,
Satoh, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K.,
Nagase, T., Nomura, N., Kikuchi, H., Masuho, Y., Yamashita, R.,
Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length
human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
JOURNAL
PUBMED
14702039
REFERENCES
AUTHORS
Ishibashi, T., Kanehori, K., Yosida, M., Watanabe, S., Ishida, S.,
Ono, Y., Hotta, T., Hirooka, S., Murakami, K., Takiguchi, S.,
Kusano, J., Watanabe, M., Fujimori, K., Tanai, H., Ishida, M.,
Yamashita, H., Chiba, Y., Sugiyama, T., Irie, R., Otsuki, T., Sato, H.,
Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, H.,
Saito, K., Nishikawa, T., Kimura, K., Matsuo, K., Nakamura, Y.,
Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Takahashi, Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahara, K., Masuho, Y., Nagai, K. and Isegai, T.
NEDO human cDNA sequencing project
Unpublished
JOURNAL
REFERENCE
AUTHORS
Isegai, T. and Yamamoto, J.
TITLES
Direct Submission
Submitted (04-JUL-2002) Takao Isegai, FLJ Project (HRI Team); 2-6-7
Kazuo-Kamatari, Kiseazu, Chiba 292-0812, Japan

COMMENT

E-mail: genomics@nti.co.jp, Tel:81-338-52-3975, Fax:81-338-52-3986 (NE) human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- and 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center; National Institute of Technology and Evaluation; Clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Source

```

source
1..1811
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone="TEST12026453"
/tissue_type="testis"
/clone_1ib="TEST12"
/note="Cloning vector: pME18FL3"
150..1580
CDS

```

ORIGIN

| | | | | |
|----------------------------|-------|---------------------|-------------|--------------|
| Query Match | 70.0% | Score 890.2; | DB 9; | length 1811; |
| Best Local Similarity | 80.9% | Pred. No. 3.2e-230; | | |
| Matches 1178; Conservative | 0; | Mismatches 3; | Indels 276; | Gaps 1; |

| | | | |
|----|-----|---|-----|
| QY | 86 | GGCGTGTTCCTTCAAAATAAAGAACTGTAAGAACTGATTACACATTAGTCATG | 145 |
| Db | 125 | GGCGTGTTCCTTCAAAATAAAGAACTGTAAGAACTGATTACACATTAGTCATG | 184 |
| QY | 146 | TGACGATGTCAACTGCTGTGCCTTCCTTTTCCCTCTGGTACTTGTCTCTGGACAA | 205 |
| Db | 185 | TGACGATGTCAACTGCTGTGCCTTCCTTTTCCCTCTGGTACTTGTCTCTGGACAA | 244 |
| QY | 206 | AACATTGCGCCTGACTGTTACGGACTTTACTGAACCTGCACATTCCTCAATGAAGTT | 265 |
| Db | 245 | AACATTGCGCCTGACTGTTACGGACTTTACTGAACCTGCACATTCCTCAATGAAGTT | 304 |
| QY | 266 | TCATACCTATGCTGTGCCACTGCTGTGTTCTCCCTTCAGAGACATAATTTGGCATCGTG | 325 |
| Db | 305 | TCATACCTATGCTGTGCCACTGCTGTGTTCTCCCTTCAGAGACATAATTTGGCATCGTG | 364 |
| QY | 326 | TTCAACAGATGTGTAACACTGTGCTCTATGGAATACGAAATGGAACAGATGCTGCAGTGAT | 385 |
| Db | 365 | TTCAACAGATGTGTAACACTGTGCTCTATGGAATACGAAATGGAACAGATGCTGCAGTGAT | 424 |
| QY | 386 | GGAACAGCGTATGTGGCAGCCCTGTGAGGGGTTTGCAGATTTTCCCAGACTCAAGTGT | 445 |
| Db | 425 | GGAACAGCGTATGTGGCAGCCCTGTGAGGGGTTTGCAGATTTTCCCAGACTCAAGTGT | 484 |
| QY | 446 | GGCATCAAGGGCAGCTGATGGAACCTGTGTTTGTGGAATGACACAGTATCAAAATTATA | 505 |
| Db | 485 | GGCATCAAGGGCAGCTGATGGAACCTGTGTTTGTGGAATGACACAGTATCAAAATTATA | 544 |
| QY | 506 | TAGATGTGTAAGTGTAAAGATGCTCCTTGGCGGACATGTGCAATTTCTCCTAATGAG | 565 |
| Db | 545 | TAGATGTGTAAGTGTAAAGATGCTCCTTGGCGGACATGTGCAATTTCTCCTAATGAG | 604 |
| QY | 566 | CTTCTTTGTCACTGGCTCTCTCATGTGTGTAATTAACATGTGGGATGATTAATAATGAGTG | 625 |
| Db | 605 | CTTCTTTGTCACTGGCTCTCTCATGTGTGTAATTAACATGTGGGATGATTAATAATGAGTG | 664 |

| | | | |
|----|------|---|------|
| QY | 626 | TCGTCATAGGAAAAACCAATGATCTTTGGAAATTACCTGCTGCAATTTTCTTACAGCC | 685 |
| Db | 665 | TCGTCATAGTGAAGAAACCACTGATCTTGGAAATTACCTGCTGCAATTTTCTTACAGCC | 724 |
| QY | 686 | AGTTTCTGATGAGAAACAGAGCTCTTCAGTTTTCGACTGGCATCATGTGTCAGATG | 745 |
| Db | 725 | AGTTTCTGATGAGAAACAGAGCTCTTCAGTTTTCGACTGGCATCATGTGTCAGATG | 784 |
| QY | 746 | CCAAAGTCAAAATTTGGAATTTGTTCTTTTACCCTATCTT----- | 784 |
| Db | 785 | CCAAAGTCAAAATTTGGAATTTGTTCTTTTACCCTATCTTAGTTTGAATTTAAATATAA | 844 |
| QY | 785 | ----- | 784 |
| Db | 845 | AAGTACACTGAGTGGGACACTGTGCTCCTGTTCTGGCTTTGCTTTTCCATGATGGCA | 904 |
| QY | 785 | ----- | 784 |
| Db | 905 | GATGCTAGTCTCAGAGGTCAGTGGATAGTGTGTCATAGTATGATAGTACTAATACTGAGAA | 964 |
| QY | 785 | ----- | 784 |
| Db | 965 | TATACTTCAACATTTGACTGAGCAACAGGTAATGTCACAACCTTGCTTTTGCACTTAA | 1024 |
| QY | 785 | ----- | 784 |
| Db | 1025 | TACCCTTTTACTTGCTCTACTGCTTCAATGACAAAACGTGACATCTGGCAATTTGACCT | 1084 |
| QY | 785 | -----ACCAAGGCGCACAGAACATCAGCTGAGCAATTTTACCGAAGATG | 829 |
| Db | 1085 | GGAACAACCTTGGCCANCCAGAGACAGAACATCAGCTGAGCAATTTTACCGAAGATG | 1144 |
| QY | 830 | GTCAGAGAGGTGCTCTCAACATGGCTTTGTGCAACAAGATTTTAAAGATCTTGTGGTAT | 889 |
| Db | 1145 | GTCAGAGAGAGGTGCTCTCAACATGGCTTTGTGCAACAAGATTTTAAAGATCTTGTGGTAT | 1204 |
| QY | 890 | TTTTCAAATGAATTAACCTGATGGAAGAAAGCTGTTGAATCTTACAAAGAAAGTCTGGC | 949 |
| Db | 1205 | TTTTCAAATGAATTAACCTGATGGAAGAAAGCTGTTGAATCTTACAAAGAAAGTCTGGC | 1264 |
| QY | 950 | TGATGATTTGAAAATTGAAATCTCTAGGACTGCGTAGTAAAGTGTGAGAAAATTGAGA | 1009 |
| Db | 1265 | TGATGATTTGAAAATTGAAATCTCTAGGACTGCGTAGTAAAGTGTGAGAAAATTGAGA | 1324 |
| QY | 1010 | GCTCAGAGCAAGGTTAAATCCCTTTCTTCAAGAAATCTGATGAATTTATATGTCCAAT | 1069 |
| Db | 1325 | GCTCAGAGCAAGGTTAAATCCCTTTCTTCAAGAAATCTGATGAATTTATATGTCCAAT | 1384 |
| QY | 1070 | AACTAGAGAACTTATGAAAATCCGGTCATGCGACTGAGATGGCTATCATATGAAAAGGA | 1129 |
| Db | 1385 | AACTAGAGAACTTATGAAAATCCGGTCATGCGACTGAGATGGCTATCATATGAAAAGGA | 1444 |
| QY | 1130 | AGCAATGGAATTTGATCAGCAAAAAGAACGTACAGTCCCATGACAAATCTTGTTCT | 1189 |
| Db | 1445 | AGCAATGGAATTTGATCAGCAAAAAGAACGTACAGTCCCATGACAAATCTTGTTCT | 1504 |
| QY | 1190 | TCCTTCAGCGGTACTTACACCAATATGAGACTCTGAAAATGGCCATCATATGATGGCTGGA | 1249 |
| Db | 1505 | TCCTTCAGCGGTACTTACACCAATATGAGACTCTGAAAATGGCCATCATATGATGGCTGGA | 1564 |
| QY | 1250 | GACACACCAAAAGTAA 1266 | |
| Db | 1565 | GACACACCAAAAGTAA 1581 | |

RESULT 7
CO491312
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS

CO491312
Sequence 23179 from Patent WO01060869
CO491312
CO491312.1 GI:41456931

1596 bp
DNA
linear

PAT 30-JAN-2004

| | |
|----------------------------|--|
| SOURCE | Homo sapiens (human) |
| ORGANISM | Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. |
| REFERENCE | 1 Schlegel R., Endege W.O. and Monahan J.E. Genes differentially expressed in human prostate cancer and their use |
| AUTHORS TITLE | |
| JOURNAL | Patent: WO 0160860-A 23179 23-AUG-2001; Millennium Predictive Medicine, Inc. (US) Location/Qualifiers |
| FEATURES | source 1..1996 /organism="Homo sapiens" /mol_type="unasigned DNA" /db_xref="taxon:9606" |
| ORIGIN | |
| Query Match | 69.1%; Score 878.8; DB 6; Length 1996; |
| Best Local Similarity | 80.7%; Pred. No. 4e-227; |
| Matches 1166; Conservative | 0; Mismatches 2; Indels 276; Gaps 1; |
| Dy | 99 CAAATTAAGAACAATGCTGAACCTGATTCAACATTAGCTGATGTGACGATCTAAC 158 |
| Dd | 4 CAAATTAAGAACAATGCTGAACCTGATTCAACATTAGCTGATGTGACGATCTAAC 63 |
| Oy | 159 TGCATGCTCTCCTTTTCCCTCTGGCTACTTGCTCTTGAGCAAAAACATTCGCTG 218 |
| Dd | 64 TGCTGTGCTCTCTCTTTCCCTCTGGCTACTTGCTCTTGAGCAAAAACATTCGCTG 123 |
| Oy | 219 TACTGTACGTGACTTACTTACTGAATGCGACATTCCTCATAGATTTCATACCTAGCT 278 |
| Dd | 124 TACTGTACGTGACTTACTTACTGAATGCGACATTCCTCATAGATTTCATACCTAGCT 183 |
| Oy | 279 GTCCACTGCTGCTGTTTTTCCCCTCAGACATATTTTGGCATGCTGTTCACACATGCT 338 |
| Dd | 184 GTCCACTGCTGCTGTTTTTCCCCTCAGACATATTTTGGCATGCTGTTCACACATGCT 243 |
| Oy | 339 ACCACTGCCATAGGAATACTGAATAATGACAGATGCTGGCAGTAGTGAACAGCTAGT 398 |
| Dd | 244 ACCACTGCCATAGGAATACTGAATAATGACAGATGCTGGCAGTAGTGAACAGCTAGT 303 |
| Oy | 399 GGCA GCCCTGTGAGGGTTTGGCAGTTTTTCCCAGACTCACGCTGTTTTGGCATCAGGGCA 458 |
| Dd | 304 GGCA GCCCTGTGAGGGTTTGGCAGTTTTTCCCAGACTCACGCTGTTTTGGCATCAGGGCA 363 |
| Oy | 459 GCTGATGGAATCTGTGTTTTTGTGGAAATGACAGATCATCAAAATATATAGATGTGTAGT 518 |
| Dd | 364 GCTGATGGAATCTGTGTTTTTGTGGAAATGACAGATCATCAAAATATATAGATGTGTAGT 423 |
| Oy | 519 GTTAAGAATGCTCCTTGGCGGCATATGSCATTTTCTCTAATGGAAGCTTTTGTGCACT 578 |
| Dd | 424 GTTAAGAATGCTCCTTGGCGGCATATGSCATTTTCTCTAATGGAAGCTTTTGTGCACT 483 |
| Oy | 579 GGCTCTCATGTGATGATTTAAACAGTGTGGAGTGAATAAATGAGGTGTCTCATAGTGA 638 |
| Dd | 484 GGCTCTCATGTGATGATTTAAACAGTGTGGAGTGAATAAATGAGGTGTCTCATAGTGA 543 |
| Oy | 639 AAAGCACAATGATCTTGAATTAAGTCTGCTGCGATTTTTTCTTACACAGCCAGTTTCTGATGA 698 |
| Dd | 544 AAAGCACAATGATCTTGAATTAAGTCTGCTGCGATTTTTTCTTACACAGCCAGTTTCTGATGA 603 |
| Oy | 699 GAACAAGGCTCTTCAAGTTTTTTCGACTGGCATCATGSGTCAGAGATTGCAAGTCAAAATT 758 |
| Dd | 604 GAACAAGGCTCTTCAAGTTTTTTCGACTGGCATCATGSGTCAGAGATTGCAAGTCAAAATT 663 |
| Oy | 759 TGAATGTTTCTTTTACCATACTT----- 784 |
| Dd | 664 TGAATGTTTCTTTTACCATACTTAAATAATAAATGAACA CTAGT 723 |
| Oy | 785 ----- 784 |
| Dd | 724 GGGCAGTGTGCTCTGTTCTGGCTTGCTTTTCCCATGATGGCAGATGCTAGTCA 783 |

| | | | | | |
|------------|---|---|------|--------|-----------------|
| OY | 785 | ----- | 784 | ----- | 784 |
| Db | 784 | GGGTCACTGGATTAAGTCTGTCAATAGTATTAAGTAACTAATACTAGAAATATACTTCACACA | 843 | | |
| OY | 785 | ----- | 784 | ----- | 784 |
| Db | 844 | TTGACTAGGACACACAGGTATATGCACAACTTGTCCTTTTGACCTAATACCTTTTACTT | 903 | | |
| OY | 785 | ----- | 784 | ----- | 784 |
| Db | 904 | GCTACTGTTCAATGACAAACAGTGACATCTGCAATTTGACCTGAAACACTTTGC | 963 | | |
| OY | 785 | --ACGAAAGGCGACAGAACATCAAGCTGAGCAATTTTACCGAAGTTGGTCAGAGAGTC | 842 | | |
| Db | 964 | CACACAAAGGGCACACAGAACATCAGCTCAGACCAATTTTCCAGAGATGGTCAGAGAGAGAT | 1023 | | |
| OY | 843 | GTCCTCAACATGGCTTTGTGACACAGATTTTAAAGATCTTGTTGATTTTCAAGATGAT | 902 | | |
| Db | 1024 | GTCCTCAACATGGCTTTGTGACACAGATTTTAAAGATCTTGTTGATTTTCAAGATGAT | 1083 | | |
| OY | 903 | AACCTTGATGGAAGAACTGTTGAATCTTACAAAGAAAGTCGTGCTGATGATTTGAAA | 962 | | |
| Db | 1084 | AACATTTGATGGAAGAACTGTTGAATCTTACAAAGAAAGTCGTGCTGATGATTTGAAA | 1143 | | |
| OY | 963 | ATTGAATCTCTAGGACTGCGGATGTAAGTAAAGTCGTGAGAAATTTGAAGCTCAGAGACCAAG | 1022 | | |
| Db | 1144 | ATTGAATCTCTAGGACTGCGGATGTAAGTAAAGTCGTGAGAAATTTGAAGCTCAGAGACCAAG | 1203 | | |
| OY | 1023 | GTTAAATCCCTTCTTCAGGAATTCCTGATGAATTTATATGTCCAATTAACCTAGAGAACTT | 1082 | | |
| Db | 1204 | GTTAAATCCCTTCTTCAGGAATTCCTGATGAATTTATATGTCCAATTAACCTAGAGAACTT | 1263 | | |
| OY | 1083 | ATGAAAGATCCGGTCATCGCATCAGATGGCTATTTCAATTAAGAAAGAAAGCAATGGAAAT | 1142 | | |
| Db | 1264 | ATGAAAGATCCGGTCATCGCATCAGATGGCTATTTCAATTAAGAAAGAAAGCAATGGAAAT | 1323 | | |
| OY | 1143 | TGATTCAGCAAAAAAGAAAGTACAAAGTCCCATGACAAATCTTGTTCTTCTTACAGCGGTA | 1202 | | |
| Db | 1324 | TGATTCAGCAAAAAAGAAAGTACAAAGTCCCATGACAAATCTTGTTCTTCTTACAGCGGTA | 1383 | | |
| OY | 1203 | CTTACACCAATATAGGACTCTGAAAAATGGCCATCAATATGATGGCTGAGACACACCAAAAG | 1262 | | |
| Db | 1384 | CTTACACCAATATAGGACTCTGAAAAATGGCCATCAATATGATGGCTGAGACACACCAAAAG | 1443 | | |
| OY | 1263 | TAAA 1266 | | | |
| Db | 1444 | TAAA 1447 | | | |
| RESULT 8 | | | | | |
| CO497179 | | | | | |
| LOCUS | CO497179 | 1996 bp | DNA | linear | PAT 30-JAN-2004 |
| DEFINITION | Sequence 29046 from Patent WO0160860. | | | | |
| ACCESSION | CO497179 | | | | |
| VERSION | CO497179.1 | GI:41462815 | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | Homo sapiens (human) | | | | |
| | Homo sapiens | | | | |
| | Eumetazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiindae; Homo. | | | | |
| REFERENCE | 1 | Schlegel, R., Endege, W. O. and Monahan, J. E. | | | |
| AUTHORS | | Genes differentially expressed in human prostate cancer and their | | | |
| TITLE | | use | | | |
| JOURNAL | | Patent: WO 0160860-A 29046 23-AUG-2001; | | | |
| FEATURES | | Millennium Predictive Medicine, Inc. (US) | | | |
| source | | location/Qualifiers | | | |
| | | 1. .1996 | | | |
| | | /organism="Homo sapiens" | | | |
| | | /mol_type="unassigned DNA" | | | |
| | | /db_xref="taxon:9606" | | | |
| ORIGIN | | | | | |

| | | | | | | |
|-----------------------|-----|---|------------------|------------|-------------|------------|
| Query Match | | 69.1% | Score 878.8 | DB 6 | Length 1996 | |
| Best Local Similarity | | 80.7% | Pred. No. 4e-227 | | | |
| Matches 1166 | | Conservative | 0 | Mismatches | 2 | Indels 276 |
| Gaps | | | | | | 1 |
| QY | 99 | CAAAATAAAGAACATGCTGTAAGAACTGATTCACACATTAGGTCATCATGTGTCAGCATGTCAAC | 158 | | | |
| DB | 4 | CAAAATAAAGAACATGCTGTAAGAACTGATTCACACATTAGGTCATCATGTGTCAGCATGTCAAC | 63 | | | |
| QY | 159 | TGCTGTGCTTCTCCTTTTCCCTCTTGGCTACTGCTCTCTTGGACAAACAATTGCTGTG | 218 | | | |
| DB | 64 | TGCTGTGCTTCTCCTTTTCCCTCTTGGCTACTGCTCTCTTGGACAAACAATTGCTGTG | 123 | | | |
| QY | 219 | TACTCGTTAGTGACTTACTGAACTGCGACATTCCTCATTTGAAGTTTCATACCTATGTG | 278 | | | |
| DB | 124 | TACTCGTTAGTGACTTACTGAACTGCGACATTCCTCATTTGAAGTTTCATACCTATGTG | 183 | | | |
| QY | 279 | GTCCACTGCTGCTGTTTCTCCCTTCAGAGACATATTTTGGCATCGTGTTCACACATAGGT | 338 | | | |
| DB | 184 | GTCCACTGCTGCTGTTTCTCCCTTCAGAGACATATTTTGGCATCGTGTTCACACATAGGT | 243 | | | |
| QY | 339 | ACCACTGCTCTATGGAATATCGAAAATGACACATGCTGGCAGTATGGAACAGCTTACT | 398 | | | |
| DB | 244 | ACCACTGCTCTATGGAATATCGAAAATGACACATGCTGGCAGTATGGAACAGCTTACT | 303 | | | |
| QY | 399 | GGCAGCCCTGAGAGGTTTGGCAGTTTCCCAAGCTCACGCTGTTTGGCACACAGGGGA | 458 | | | |
| DB | 304 | GGCAGCCCTGAGAGGTTTGGCAGTTTCCCAAGCTCACGCTGTTTGGCACACAGGGGA | 363 | | | |
| QY | 459 | GCTATGGAACGTGTGTTTGTGGAAATGACACGTATACAAATTAATATGATGTGTAGT | 518 | | | |
| DB | 364 | GCTATGGAACGTGTGTTTGTGGAAATGACACGTATACAAATTAATATGATGTGTAGT | 423 | | | |
| QY | 519 | GTTAAAGATGCTCCTTGGCGGCATGTGCATTTTCTCTAATGAGAGCTTCTTGTCACT | 578 | | | |
| DB | 424 | GTTAAAGATGCTCCTTGGCGGCATGTGCATTTTCTCTAATGAGAGCTTCTTGTCACT | 483 | | | |
| QY | 579 | GGCTCTCATGTGTGATTTTAAACATGTGGGATGATTAATAAGTGTCTCATATGTGA | 638 | | | |
| DB | 484 | GGCTCTCATGTGTGATTTTAAACATGTGGGATGATTAATAAGTGTCTCATATGTGA | 543 | | | |
| QY | 639 | AAAGCACATGATCTTGGAAATTAACCTGCTGGCATTTTCTTCAACGCCAGTTCTGATGGA | 698 | | | |
| DB | 544 | AAAGCACATGATCTTGGAAATTAACCTGCTGGCATTTTCTTCAACGCCAGTTCTGATGGA | 603 | | | |
| QY | 699 | GAACAAGTCTTCAAGTTTTCATGTCATGTCATCATGTGGTCAGAGATTGCCAATGCAAAAT | 758 | | | |
| DB | 604 | GAACAAGTCTTCAAGTTTTCATGTCATGTCATCATGTGGTCAGAGATTGCCAATGCAAAAT | 663 | | | |
| QY | 759 | TGATTTGTTTCTTTTACCCTATCTT----- | 784 | | | |
| DB | 664 | TGATTTGTTTCTTTTACCCTATCTT----- | 723 | | | |
| QY | 785 | ----- | 784 | | | |
| DB | 724 | GGGCACTGTGCTCCTGTTCTGGCTTGTGCTTTTCCATGATGGCGAGTGTACTCTCA | 783 | | | |
| QY | 785 | ----- | 784 | | | |
| DB | 784 | GGGTCAAGTGAATAGTCTGTCACTACTATATGATACTAATACTGAGAAATATCTTACACACA | 843 | | | |
| QY | 785 | ----- | 784 | | | |
| DB | 844 | TTGACTACGACACACAGGATATGTCACACTTGTCTTTTGACCTAATACCTTTACTTT | 903 | | | |
| QY | 785 | ----- | 784 | | | |
| DB | 904 | GCTACTGTGTTCAATGACAAAACAATGAACATCTGGCAATTTGACCTGAAAACACTTTGC | 963 | | | |
| QY | 785 | --ACCAAGGCGACAGAACATCAGCTGGAACAATTTTACGAAAGATTGTGCAGAGAGTGC | 842 | | | |
| DB | 964 | CAAGCAAGGCGCACAGAACATCAGCTGGAACAATTTTACGAAAGATTGTGCAGAGAGAT | 1022 | | | |
| QY | 843 | GTCTCAACATGAGCTTTTGTGCAACAAGATTTTAAAGATCTTGTGTATTTTCAAGATGAAT | 902 | | | |

| Db | 1024 | GTCTCAACATGCGCTTTGTGCAACAAGATTATAAAGATCTTGTTGGTATTTTCAAGATGAAT | 1083 |
|-------------------------|--|--|--------------------|
| Qy | 903 | AACATGATGAGAAAGAACTGTTGAATCTTACAAAGAAAGTCTGCGTATGATTTGAA | 962 |
| Db | 1084 | AACATTTGAGAAAGAACTGTTGAATCTTACAAAGAAAGTCTGCGTATGATTTGAA | 1143 |
| Qy | 963 | ATTGAATCTTAGAGACCTGCGTAACTGTAAGTGTGAGGAAATTTGAAGAGCTTCAGAACAG | 1022 |
| Db | 1144 | ATTGAATCTTAGAGACCTGCGTAACTGTAAGTGTGAGGAAATTTGAAGAGCTTCAGAACAG | 1203 |
| Qy | 1023 | GTTAAATCCCTTTCTTACAGGAATCTCGATGAATTTATATGTCCAAATACATAGAGAACTT | 1082 |
| Db | 1204 | GTTAAATCCCTTTCTTACAGGAATCTCGATGAATTTATATGTCCAAATACATAGAGAACTT | 1263 |
| Qy | 1083 | ATGAAGATCCGCTCATGCGATCGATCAGATGGCTATTATATGAGAAAGAAAGCAATGAAAT | 1142 |
| Db | 1264 | ATGAAGATCCGCTCATGCGATCGATCAGATGGCTATTATATGAGAAAGAAAGCAATGAAAT | 1323 |
| Qy | 1143 | TGATTCAGCAAAAAAGAAAGCTACAGATGCCATGACAAATCTTGTCTTCCCTTACGGGTA | 1202 |
| Db | 1324 | TGATTCAGCAAAAAAGAAAGCTACAGATGCCATGACAAATCTTGTCTTCCCTTACGGGTA | 1383 |
| Qy | 1203 | CTTACACCAATATAGGACTCTGAAATATGGCATCAATGATGGCTGAGACACACCAAAAG | 1262 |
| Db | 1384 | CTTACACCAATATAGGACTCTGAAATATGGCATCAATGATGGCTGAGACACACCAAAAG | 1443 |
| Qy | 1263 | TAAA | 1266 |
| Db | 1444 | TAAA | 1447 |
| RESULT 9 | | | |
| AX077675 | | | |
| LOCUS | AX077675 | 1291 bp | DNA |
| DEFINITION | Sequence 130 from Patent WO0105970. | | linear |
| ACCESSION | AX077675 | | |
| VERSION | AX077675.1 | GI:13122057 | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | Homo sapiens (human) | | |
| REFERENCE | | | |
| AUTHORS | 1 | | |
| TITLE | Yue, H., Tang, Y. T., Bandman, O., Hillman, J. L., Lai, P., Hu, Young, J., | | |
| JOURNAL | Reddy, R., Yang, J., Baughn, M. R., Lu, D. A., Azimzai, Y. and | | |
| FEATURES | Peterson, C. | | |
| source | Genp-binding protein associated factors | | |
| | Patent: WO 0105970-A 10 25-JAN-2001; | | |
| | Incyte Genomics, Inc. (US) | | |
| | Location/Qualifiers | | |
| | 1..1291 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="unassigned DNA" | | |
| | /db_xref="taxon:9606" | | |
| | /note="Incyte ID No: 4546403CB1" | | |
| ORIGIN | | | |
| Query Match | 37.6% | Score 478.6; | DB 6; Length 1291; |
| Best Local Similarity | 96.2% | Pred. No. 1.6e-118; | |
| Match 484; Conservative | 0; | Mismatches 9; | Indels 0; Gaps 0; |
| Qy | 774 | ACCATATCTTACGAGGCGACAGAACTACGCTGAGCAATTTACGAAGATTGTGCA | 833 |
| Db | 547 | ACACTTTCAGAGAGAGAGACAGCAACATACGCTGAAGCAATTTACGAAGATTGTGCA | 606 |
| Qy | 834 | GAGAGGTCGTCTCAACATGGCTTTGTGCAACAAGTTTAAAGATCTTGTGGTATTTTC | 893 |
| Db | 607 | GAGAGAGATGTCACATGGCTTTGTGCAACAAGTTTAAAGATCTTGTGGTATTTTC | 666 |
| Qy | 894 | AAGATGAATACATTTATGAGAAAGAACTGTGTAATCTTACAAAGAAAGTCTGGCTGAT | 953 |
| Db | 667 | AAGATGAATACATTTATGAGAAAGAACTGTGTAATCTTACAAAGAAAGTCTGGCTGAT | 726 |

QY 954 GATTGAAATTTGATCTCTAGAGCTGCGTAGTAAAGTCTGAGAAATTTGAAGAGCTC 1013
DB 727 GATTGAAATTTGATCTCTAGAGCTGCGTAGTAAAGTCTGAGAAATTTGAAGAGCTC 786
QY 1014 AGGACCAAGGTTAAATCCCTTTCTTTCAGAAATTCCTGATGAATTTATATGCTCAATTAAT 1073
DB 787 AGGACCAAGGTTAAATCCCTTTCTTTCAGAAATTCCTGATGAATTTATATGCTCAATTAAT 846
QY 1074 AGGAACTTATGAAAGTCCGCTGATCGATCGATGAGTCTTATTCATATGAAAGGAAGCA 1133
DB 847 AGGAACTTATGAAAGTCCGCTGATCGATCGATGAGTCTTATTCATATGAAAGGAAGCA 906
QY 1134 ATGGAATTTGATTCAGCAAAAAGAAACGTACAGTCCATGACAAATCTTGTCTTCTCT 1193
DB 907 ATGGAATTTGATTCAGCAAAAAGAAACGTACAGTCCATGACAAATCTTGTCTTCTCT 966
QY 1194 TCAGCGCTACTTCACCAATATAGACTCTGAAAATGGCCATCAATGATGCTGAGAGCA 1253
DB 967 TCAGCGCTACTTCACCAATATAGACTCTGAAAATGGCCATCAATGATGCTGAGAGCA 1026
QY 1254 CACCAAAAGTAAA 1266
DB 1027 CACCAAAAGTAAA 1039

RESULT 10
LOCUS BC061948 2130 bp mRNA 1linear ROD 19-NOV-2003
DEFINITION Mus musculus RIKEN cDNA 2610014F08 gene, mRNA (cDNA clone MGC:65547
IMAGE:6312489), complete cds.
ACCESSION BC061948
VERSION BC061948.1 GI:38303926
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE
AUTHORS Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.L., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton E., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schectz T.E., Brownstein M.J., Uedlin T.B., Tomihata S.,
Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullaney S.J., Bosak S.A., McGowan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Vallblat D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettner M., Madan A., Young A.C., Shcherbina Y.,
Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smallog D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

TITLE
JOURNAL MEDLINE
PUBMED 22388257
REFERENCE 2 (bases 1 to 2130)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (10-NOV-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIA-MGC Project URL: <http://mgc.ncl.nih.gov>
COMMENT Contact: MGC help desk
Email: cgads-f@mail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.

CDNA Library Preparation: ResGen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@ncl.nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dierich, N.L., Grante, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
Maduro, Q.L., Mastello, C., Masker, B., Mastrian, S.D., McCreasey, J.C.,
McDowell, J., Pearson, R., Stantrop, S., Thomas, P.J., Touchman, J.W.,
Taurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRM Plate: 127 Row: 1 Column: 1
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.

FEATURES
source

gene

CDS

misc_feature

misc_feature

misc_feature

1..2130
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:65547 IMAGE:6312489"
/tissue_type="olfactory epithelium, neonatal mouse,
CS7BL/6"
/clone_11b="NIH_MGC_129"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"
1..2130
/gene="2610014F08R1k"
/note="Synonym: 1700048B19R1k"
/db_xref="LOCUSID:72137"
/db_xref="MGI:1919387"
208..1632
/gene="2610014F08R1k"
/translation="MVRLLTADGDVSCCAFAALLATGSLDKTRLYSLSPAE
LSPSPKRTTVAHCCSPSPGSHVLAAGSTDDTTLVMSHSHRTITVLEQPGSPYRV
CCFSPSAYLADGADGSLALNNAQTYKYRCGSVDSLVACAFSPDGLPVTSSG
GDTVWDDMRKCHSEKADIDITCCFSSQPSGEGSLQSLACGQDCBTKLMAV
TTRTVGPELTKYKSLTSGHCAVLACAFSDHDKMLAGSDVSVIHHGIPQSVATL
TQTRVYVTCAPAPNTLLATSGMKTVMVQFDELTGCGSMNPLKHTFEMSEED
VSVMLRAQGLDELNGLFRANNIDGKELHLTRESLAGDKIESLGRKVSRLSIELR
AKMDLSISGIPPEFICPTIRELMKQPVLASDGYSTEREMESHIKKRTSPMTIAL
PSLVLPNTLWAINRWLETHK"
217..1131
/gene="2610014F08R1k"
/note="WP40: Region: WD40 domain, found in a number of
eukaryotic proteins that cover a wide variety of functions
including adaptor/regulatory modules in signal
transduction, pre-mRNA processing and cytoskeleton
assembly"
/db_xref="CD:cd00200"
1192..1183
/gene="2610014F08R1k"
/note="SAM: Region: Sterile alpha motif."
/db_xref="CD:cd00166"
1408..1629
/gene="2610014F08R1k"
/note="U-box: Region: U-box domain. This domain is related
to the Ring finger Pfam00097 but lacks the zinc binding
residues"
/db_xref="CD:pfam04564"

ORIGIN

| | | | | | | | |
|-----------------------|-------|--------------|---------------|------------|------|--------|-------|
| Query Match | 37.3% | Score | 474.6; | DB | 10; | Length | 2130; |
| Best Local Similarity | 62.8% | Pred | No. 2.1e-117; | | | | |
| Matches | 955; | Conservative | 0; | Mismatches | 289; | Indels | 277; |
| | | | | | | Gaps | 3 |

| | | | |
|----|------|---|------|
| QY | 19 | CGCGCAGCTGATACCCGACGCCGCCGTGGGACCTTGAAGGGGGATATCCCGCGGCCCGAC | 78 |
| Db | 117 | CTGCACCCCGGAGCCGCGCAGGGGAGAGCCGCCCGAGGGGGGACTTCACTCGGGGTG | 176 |
| QY | 79 | TCCTGCAGGCTGTGTTTTCTTGAAAATAAGAACATGTGTGAAAATGATTCACATTAAGCTG | 138 |
| Db | 177 | CGAGGGAAAGCTG-TTAACTTTGTGTGAGGAACAATGTATGAGGTGATTCACAGCTGGCTG | 235 |
| QY | 139 | ATCATGTGTGAAGATGTCAACTGTGTGTGCTTCTCTCTTTCCCTCTTGGCTACTTGTCTCT | 198 |
| Db | 236 | ATCACGGTGAAGAGCTCAGCTGTGCGCTTCTCGGCTGCCCTCCTGTGCCACCTGTCTCT | 295 |
| QY | 199 | TGACAAATAAATTCGCTGTACTGTCTGTACGAGACTTACTGAACTGGCAATTCCTCCAT | 258 |
| Db | 296 | TGGACAAACCATTCGCGCTGTACTCTCCATGATGACTTTGCTGMACTGCGATCTCCCGC | 355 |
| QY | 259 | TGAAGTTTCATTAACCTATGTCTGTCACTGTCTGTCTTCTCCCTTCAGAGACATAATTTGG | 318 |
| Db | 356 | TGAAGTTTCACACTATGTCTGTCCACTGTCTGTGTTTCTACCCCTCAGAGACAGCTTTAG | 415 |
| QY | 319 | CATGTGTTCACAGATGTGTACAATGTCTTCACTGTCTCTATGGAAATCTGAAATAATGGAATGCTGG | 378 |
| Db | 416 | CATGTGTGTGCACAGACGGGACACAGGTGTGTGAGTGTGCACAGCGGACACACCCCTGA | 475 |
| QY | 379 | CAGTATATGGAACACGCTTGTGTGGCAGCCCTGTGAGGGATTGTGCAGTTTTCCCACTTCA | 438 |
| Db | 476 | CCGTTGTGAGACAGCGGGGTGGCAGCCCTGTGTGCGGTCTGTGTGCTTTTCCCAACTCTG | 535 |
| QY | 439 | CGTGTTTGGCATCAGGGGCACTGTATGGAATCTGTGTTTGTGGAATGCACAGTCATACA | 498 |
| Db | 536 | CCTACCTTAGCCCTCAGGGGCTGCCGATGGATCCATTTGCTTGTGGAATGCACAGACATACA | 595 |
| QY | 499 | AATTAATATAGATGTGTAGTGTATAAGATGCTCTCTGGGGGCATATGSCATTTTCTCCTA | 558 |
| Db | 596 | AACATAATAGGTGTGTATGTATCAAGATATCTCAATGTGTGACCTGTGTGCTTCTCCG | 655 |
| QY | 559 | ATGGAAGCTTCTTTGTCACTGGCTCTTCATGTGTGATTTAAACAGTGTGGATGATAAA | 618 |
| Db | 656 | ATGGAAGGCTCTTTGTGTCACTGGCTCTCTGGCGGGGACTTGACAGTGTGGGATGACAGA | 715 |
| QY | 619 | TGAGGTGTCTGCATATGTGAAAAAGCACATGATCTTGGAATTAAGCTGTGCGATTTTCTT | 678 |
| Db | 716 | TGAGGTGTCTGCACAGCGCAGAGGGGACAGATCTCGGGATCACTGTGCGAGCTTTTCTT | 775 |
| QY | 679 | CACAGCCAGTTTCTGATGTGAGAACAGAGTCTTCAAGTTTTCGATGCGGCATATGTGTC | 738 |
| Db | 776 | CACAGCCCTCTCTGGCGGAG---AAGGCTCTCAGTCTTACACAGTTGGCGTCATATGTGTC | 832 |
| QY | 739 | AGGATTTGCCAAGTCAAAAATTTGATATGTTCTTTTACCACATATAGG----- | 786 |
| Db | 833 | AAGACTGTGAATCAAACTCTGGGCTGTACTATTAACCGTGCTGTAGGCTTTGAATTA | 892 |
| QY | 787 | ----- | 786 |
| Db | 893 | AATATTAAGCACACTAATGTGGGCACTGGCCCTGTCTTGGCCTGTGCTTTTTCACATG | 952 |
| QY | 787 | ----- | 786 |
| Db | 953 | ATGGAAGATGCTTGATCGGGGGTCAGTGTAAATCTGTATCATATACATGTATCGGCC | 1012 |
| QY | 787 | ----- | 786 |
| Db | 1013 | CTCAGAGTGTCTACACACGCTGACTGACATACACAGTATGTTAACGACTTGTGGGTTTG | 1072 |
| QY | 787 | ----- | 786 |
| Db | 1073 | CACCCAAACACTCTTACTGTCTACTGTGTCAATGTGACAGACAGTGAACATTTGGCAGT | 1132 |

| QY | 787 | -----CAAGGGCA CAGAACTATGACGTGAACCAATTTACCGAAG | 825 |
|------------|--|--|------|
| Db | 1133 | TTGACCTGGAAACACCTTGGCCAMGAAAGCATGAACGCCGCTGAAACATTTCACTGAAG | 1192 |
| QY | 826 | ATTGGTCAGAGAGAGTGTCTCAACATGCGCTTTGTGCACAAGATTTTAAAGATCTTGTG | 885 |
| Db | 1193 | AATGTCAGAGAGAGATGTCTCCGTGTGTGCTGTGCTCAAGGCTTGAAAGACTCGTGG | 1252 |
| QY | 886 | GTATTTTCAAGATGAATPAACATTGATGGAAAAGAACTGTTGATCTTACAAAGAAAGTC | 945 |
| Db | 1253 | GTATTTTCAGGGCAAAACACATGCATGTGGAAAGAACTATTGTGCATCTCAAAAGAAAGTC | 1312 |
| QY | 946 | TGCGTGATGATTTGAAATTTGAATCTTAGAGACTGCGTAGTAAAGTCGTGAGAAATTTG | 1005 |
| Db | 1313 | TGGCTGGTGATTTGAAATGCAATCTTAGGCGTGCACACCAAGTCTTGAGAGATTTTG | 1372 |
| QY | 1006 | AAGAGCTCAGAGACCAAGGTTAAATCCCTTTCTTCAGAAATTCCTGATGAATTTATATGTC | 1065 |
| Db | 1373 | AAGAGCTCAGAGGCCAATATGATTTCCCTCTCTTCGGAATCCCTGACGAGTTCATCTGCC | 1432 |
| QY | 1066 | CAATPACTAGAGAACTTATGAAAGATCCGTCATGCGATCAAGATGCTTATCATATGAA | 1125 |
| Db | 1433 | CAATPAACAGAGAACTCATAGAAAGACCCGCTCATGCGATCAGATGGCTACTCTTACAGAA | 1492 |
| QY | 1126 | AGGAAGCAATGGAATAATTTGATGTCAGCAAAAAAGAAAGTACAAAGTCCCATGACAAATCTTG | 1185 |
| Db | 1493 | GAGAAAGCAATGGAAGACTGATTCACAAAGAAAGCTTACAGAGCCCATATGCAAAATTTGG | 1552 |
| QY | 1186 | TTCTTCTTCACGCGGTACTTACACCAATAGAGCTGTGAAATATGGCCATCATATGATGCG | 1245 |
| Db | 1553 | CTCTCCCTTCACTGTGTACTGACCCCAAAAGGACACTGAAGATGGCCATCAACGATGCG | 1612 |
| QY | 1246 | TGGAGACACACCAAAAGTAA | 1266 |
| Db | 1613 | TGGAGACGACACGAAAGTGAA | 1633 |
| RESULT 11 | | | |
| LOCUS | BC050792 | | |
| DEFINITION | Mus musculus RIKEN cDNA 2610014F08 gene, mRNA (cDNA clone | | |
| | IMAGE:6742025), partial cds. | | |
| ACCESSION | BC050792 | | |
| VERSION | BC050792.1 | | |
| KEYWORDS | GI:30048136 | | |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Mus musculus | | |
| REFERENCE | Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognath; Muridae; Murinae; Mus. | | |
| AUTHORS | (bases 1 to 1547) | | |
| | Strauberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shennen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heitel,F., Diatchenko,L., Marusina,K., Farmer,A.G., Rubin,G.M., Hong,L., Scapleton,M., Soares,M.B., Bonaldi,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.U., Ustin,T.B., Tenisyski,S., Carninci,P., Prange,C., Kaha,S.S., Loggellano,N.A., Peters,G.J., Aramzon,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Woley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hultyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Dittieff,D.Y.S., Krzywinski,M.I., Skalski,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A. | | |
| TITLE | Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences | | |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002) | | |
| MEDLINE | 2238257 | | |

| Query Match | 33.8% | Score 430 | DB 9 | Length 147021 |
|-----------------------|---|--|---------------|-----------------|
| Best Local Similarity | 96.7% | Pred. No. 4.1e-105 | | |
| Matches 439 | Conservative | 0 | Mismatches 15 | Indels 0 |
| | | | Gaps | 0 |
| QY | 82 | TCAGAGCTGTTTCTTCATTAATTAAGAACATGATGAACTGATTCACACATTAGTATC | 141 | |
| Db | 106602 | TTCAAGGCTGTTTCTTCATTAATTAAGAACATGATGAACTGATTCACACATTAGTATC | 106543 | |
| QY | 142 | ATGTGACGATGTCACATGCTGCTGCTTCTCTCTTTTCCCTCTTGGCTACTGCTCTTGG | 201 | |
| Db | 106542 | ATGTGACGATGTCACATGCTGCTGCTTCTCTCTTTTCCCTCTTGGCTACTGCTCTTGG | 106483 | |
| QY | 202 | ACAAAACAAATTCGCGCTGTAATCTGTTAAACGTAACCTTAATGAACTGCAACATTCCTCAATTA | 261 | |
| Db | 106482 | ACAAAACAAATTCGCGCTGTAATCTGTTAAACGTAACCTTAATGAACTGCAACATTCCTCAATTA | 106423 | |
| QY | 262 | AGTTTCATACCTATGCTGTCACACTGCTGCTGTTTCTCCCTTCAGAGCAATATTTGGCAT | 321 | |
| Db | 106422 | AGTTTCATACCTATGCTGTCACACTGCTGCTGTTTCTCCCTTCAGAGCAATATTTGGCAT | 106353 | |
| QY | 322 | CGTGTTCACACAGATGTAACCACTGTCCTTAATGTAATCTGAATAATGGAATGACAGATCTGGCAG | 381 | |
| Db | 106362 | CGTGTTCACACAGATGTAACCACTGTCCTTAATGTAATCTGAATAATGGAATGACAGATCTGGCAG | 106303 | |
| QY | 382 | TGATGGAACACAGCTATGTCGACGCGCTGTGAAGGTTTGGCCAGTTTCCCAAGATCCACAT | 441 | |
| Db | 106302 | TGATGGAACACAGCTATGTCGACGCGCTGTGAAGGTTTGGCCAGTTTCCCAAGATCCACAT | 106243 | |
| QY | 442 | GTTTGGCATCAGGGGACGCTGATGGAACATGCTGGTTTGTGGAATGACAGTCAATCAAT | 501 | |
| Db | 106242 | GTTTGGCATCAGGGGACGCTGATGGAACATGCTGGTTTGTGGAATGACAGTCAATCAAT | 106183 | |
| QY | 502 | TATATAGATGTGTAGTGTAAAGATGCTCTCTT | 535 | |
| Db | 106182 | TATATAGATGTGTAGTGTAAATTCCTCTCTGTT | 106149 | |
| RESULT 14 | | | | |
| AC008277/c | 177089 bp | DNA | linear | PRI 29-OCT-2002 |
| LOCUS | Homo sapiens BAC clone RP11-311B14 from 2, complete sequence. | | | |
| ACCESSION | AC008277 | | | |
| VERSION | AC008277.4 | GI:10716648 | | |
| KEYWORDS | HTG. | | | |
| SOURCE | Homo sapiens (human) | | | |
| ORGANISM | Homo sapiens | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | |
| REFERENCE | 1 (bases 1 to 177089) | | | |
| REFERENCE | Sulston, J.E. and Waterston, R. | | | |
| REFERENCE | Toward a complete human genome sequence | | | |
| REFERENCE | Genome Res. 8 (11), 1097-1108 (1998) | | | |
| REFERENCE | 99063792 | | | |
| REFERENCE | 9847074 | | | |
| REFERENCE | 2 (bases 1 to 177089) | | | |
| REFERENCE | Kyung, K., Maupin, R. and Hawking, M. | | | |
| REFERENCE | The sequence of Homo sapiens BAC clone RP11-311B14 | | | |
| REFERENCE | Unpublished (2001) | | | |
| REFERENCE | 3 (bases 1 to 177089) | | | |
| REFERENCE | Waterston, R.H. | | | |
| REFERENCE | Direct Submission | | | |
| REFERENCE | Submitted (31-JUL-1999) Genome Sequencing Center, Washington | | | |
| REFERENCE | University School of Medicine, 4444 Forest Park Parkway, St. Louis, | | | |
| REFERENCE | MO 63108, USA | | | |
| REFERENCE | 4 (bases 1 to 177089) | | | |
| REFERENCE | Waterston, R.H. | | | |
| REFERENCE | Direct Submission | | | |
| REFERENCE | Submitted (27-APR-2000) Genome Sequencing Center, Washington | | | |
| REFERENCE | University School of Medicine, 4444 Forest Park Parkway, St. Louis, | | | |
| REFERENCE | MO 63108, USA | | | |
| REFERENCE | 5 (bases 1 to 177089) | | | |
| REFERENCE | Waterston, R. | | | |
| REFERENCE | Direct Submission | | | |

| | | |
|-----------|---|--|
| JOURNAL | Submitted (07-OCT-2000) | Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA |
| REFERENCE | 6 (bases 1 to 177089) | |
| AUTHORS | Waterston, R.H. | |
| TITLE | Direct Submission | |
| JOURNAL | Submitted (25-MAR-2001) | Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA |
| REFERENCE | 7 (bases 1 to 177089) | |
| AUTHORS | Waterston, R. | |
| TITLE | Direct Submission | |
| JOURNAL | Submitted (29-OCT-2002) | Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA |
| COMMENT | On Oct 7, 2000 this sequence version replaced gi:7656634. | |

Center: Washington University Genome Sequencing Center
 Center code: WMGSC
 Web site: <http://genome.wustl.edu/gsc>
 Contact: baptiste@wustl.edu
 Summary Statistics
 Center project name: H_NH0311B14

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality ≥ 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:
The RpCt-II human BAC library was made from the blood of one male donor, as described by Geesagawa, K., Moon, P.Y., Zhao, B., Frensen, E., Tateno, M., Caranese, J.J. and de Jong, P.J. (1998). An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pletier de Jong and coworkers at <http://www.chori.org>
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the right is RP11-542H1, 200 base pair overlap; the clone sequenced to the left is RP11-292A10, 200 base pair overlap. Actual start of this clone is at base position 1 of RP11-311B14; actual end is at base position 23376 of RP11-542H1.

RP11-311B14 contains a transposon in the growth of the clone which has not been included in the submitted sequence. The transposon would insert after base position 212.

RP11-311B14 contains sequence from base position 23113 to 25636 which is not represented in the full sequence of it's neighbor RP11-292A10.

```
FEATURES
source
location/Qualifiers
1..177089
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="2"
/map="2"
/clone="RP11-311B14"
```

/clone_lib="RPCT-11"
repeat_region 65. .146
/rpt_family="L1"
repeat_region 185. .596
/rpt_family="L1"
repeat_region 728. .945
/rpt_family="L1"
repeat_region 959. .1171
/rpt_family="L1"
repeat_region 1365. .1969
/rpt_family="L1"
repeat_region 2185. .2293
/note="similar to EST AL040518 (NID:g5409467)"
repeat_region 2484. .3017
/rpt_family="ERVX"
repeat_region 3658. .3694
/rpt_family="L2"
repeat_region 3915. .4216
/rpt_family="Alu"
repeat_region 4720. .4861
/rpt_family="Alu"
repeat_region 5397. .5708
/rpt_family="L1"
repeat_region 6044. .6136
/note="similar to EST AL040518 (NID:g5409467)"
misc_feature 6554. .6851
/rpt_family="Alu"
repeat_region 7237. .7513
/rpt_family="Alu"
repeat_region 8885. .9148
/rpt_family="L2"
repeat_region 9369. .9554
/rpt_family="MERL_type"
misc_feature 10258. .10443
/note="similar to EST AA118718 (NID:g1676210) mp57c07.r1"
misc_feature 10259. .10452
/note="similar to EST AL040518 (NID:g5409467)"
misc_feature 10353. .10443
/note="similar to EST AA118718 (NID:g1676210) mp57c07.r1"
repeat_region 10908. .11207
/rpt_family="Alu"
repeat_region 11747. .11904
/rpt_family="Alu"
repeat_region 12188. .12490
/rpt_family="Alu"
repeat_region 12696. .12988
/rpt_family="Alu"
repeat_region 13170. .13582
/note="similar to EST AA1189142 (NID:g3740351) qd04a04.x1"
misc_feature 13170. .13411
/note="similar to EST AA118718 (NID:g1676210) mp57c07.r1"
misc_feature 13370. .13429
/note="similar to EST AA501581 (NID:g2236548) ne97g11.g1"
misc_feature 13284. .13544
/note="similar to EST AA482240 (NID:g7052346)"
repeat_region 13717. .13894
/rpt_family="MERL_type"
repeat_region 13913. .14066
/rpt_family="Alu"
repeat_region 14178. .14303
/rpt_family="MERL_type"
repeat_region 14308. .14608
/rpt_family="Alu"
repeat_region 15230. .15355
/rpt_family="L2"
repeat_region 15545. .15679
/rpt_family="L2"
repeat_region 15924. .16201
/rpt_family="L2"
repeat_region 16595. .17554
/rpt_family="Alu"
misc_feature 18311. .18637
/note="CpG_island (4GC=68.2, o/e=0.80, #CpGs=95)"
misc_feature 18311. .18637
/note="similar to EST AW170043 (NID:g6401568) xj36b07.x1"

repeat_region 18432. .18642
/rpt_family="L1"
repeat_region 18673. .18731
/rpt_family="L1"
repeat_region 18732. .19036
/rpt_family="Alu"
repeat_region 19037. .19162
/rpt_family="L1"
repeat_region 19206. .19463
/rpt_family="Alu"
repeat_region 19869. .20164
/rpt_family="Alu"
repeat_region 20283. .20576
/rpt_family="Alu"
repeat_region 20614. .20759
/rpt_family="L1"
repeat_region 21115. .21187
/rpt_family="L1"
repeat_region 21800. .22026
/rpt_family="Alu"

Query Match 33.8%; Score 430; DB 9; Length 177089;

Best Local Similarity 96.7%; Pred. No. 4.2e-105; Indels 0; Gaps 0;

Matches 439; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 82 TGCAGGCTGTTTCTTCAATTAAGACATGATGTAACATGATTCACATTAAGTATC 141
DB 13596 TTCAGGCTGTTTCTTCAATTAAGACATGATGTAACATGATTCACATTAAGTATC 13537

QY 142 ATGCTACGATGTCACATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 201
DB 13536 ATGCTACGATGTCACATGCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 13477

QY 202 ACAAAACAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261
DB 13476 ACAAAACAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13417

QY 262 AGTTTCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
DB 13416 AGTTTCATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 13357

QY 322 CGTTTCACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 381
DB 13356 CGTTTCACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 13297

QY 382 TGATGACAGCCTAGTGGCAGCCTGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 441
DB 13296 TGATGACAGCCTAGTGGCAGCCTGATGAGGATGAGGATGAGGATGAGGATGAGGATGAG 13237

QY 442 GTTTGCATCAGGCGCAGCTGATGAACTGTGTTTGTGGAATGACAGTCAATCAAT 501
DB 13236 GTTTGCATCAGGCGCAGCTGATGAACTGTGTTTGTGGAATGACAGTCAATCAAT 13177

QY 502 TATATGATGTGTGATGTTAAAGATGCTCCTT 535
DB 13176 TATATGATGTGTGATGTTAAAGATGCTCCTT 13143

Top 100

RESULT 15
CQ473162/c 446 bp DNA linear PAT 30 JUN-2004
LOCUS CQ473162
DEFINITION Sequence 5029 from Patent WO0160860
ACCESSION CQ473162
VERSION CQ473162.1 GI:41438781
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.
REFERENCE 1 Schlegel, R., Endege, W.O. and Monahan, J.E.
AUTHORS Genes differentially expressed in human prostate cancer and their
TITLE use

JOURNAL

Patent: WO 0160860-A 5029 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)

FEATURES

SOURCE

Location/Qualifiers
1..446/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match

26.3%; Score 335; DB 6; Length 446;

Best Local Similarity 96.3%; Pred. No. 1.4e-79;

Matches 365; Conservative 0; Mismatches 10; Indels 4; Gaps 2;

```
QY 412 GGGTTTGCAGGTTTCCCGACACTCCACGTTG--TTTGGCATCAGGGGACGCTGATGAA 468
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 445 GGTTTGCAGTTTTCCTCCAGACTCCACAGCTGTTTGGCATCAGGGGACGCTGATGAA 386
QY 469 CTGTGTTTGTGGAATGCACAGTCATACAAATTATATAGATGTGTAGTTTAAGATG 528
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 385 CTGTGTTTGTGGAATGCACAGTCATACAAATTATATAGATGTGTAGTTTAAGATG 326
QY 529 GCTCCTTGGCGGCATGTGCATTTTCTCTAATGGAAGCTTTGTCACTGGCTCTCAT 588
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 325 GCTCCTTGGCGGCATGTGCATTTTCTCTAATGGAAGCTTTGTCACTGGCTCTCAT 266
QY 589 GTGTGATTTTAACAGTGTGGGATGATAAAATGAGGTGTGCTGATAGTG-AAAAGCAAT 647
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 265 GTGTGATTTTAACAGTGTGGGATGATAAAATGAGGTGTGCTGATAGTGAAAAGCAAT 206
QY 648 GATCTTGAATTAACCTGCTGCATTTTCTTCAAGCCAGTTTCTGATGAGAACAAAGT 707
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 205 GATCTTGAATTAACCTGCTGCATTTTCTTCAAGCCAGTTTCTGATGAGAACAAAGT 146
QY 708 CTTCAGTTTTTTCAGTGGCATCATGTGTGAGGATTGCCAAGTCAAAATTGGATTGTT 767
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 145 CTTCAGTTTTTTCAGTGGCATCATGTGTGAGGATTGCCAAGTCAAAATTGGATTGTT 86
QY 768 TCTTTTACCCATATCTTAG 786
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 85 TCTTTTACCCATATCTTAG 67
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: February 5, 2005, 07:01:54
Job time : 5676 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: February 5, 2005, 13:45:37 ; Search time 5684 Seconds
(without alignments)
3273.540 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVKILHTLADHGDVNCAP.....LTPNRTKMAINWLETHK 384

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 2805242

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame-p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10077111/runac_04022005_121141_7181/app.query.fasta_1.583
-DB=GenEmbl -QFMT=faetap -SUFPTX=p2n21m200.rge -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bites -START=1 -END=-1 -MATRIX=bl0sum62 -TRANS=human40.cdl
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MTN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptlo -NOBM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200
-USER=US10077111.QCGN 1.14200 @runac_04022005_121141_7181 -NCPUB=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Length | DB ID | Description |
|------------|-------|--------------|-------|-------------------|
| C 1 | 300 | 14.7 | 180 | CQ110706 |
| C 2 | 300 | 14.7 | 180 | CQ149443 Sequence |
| C 3 | 300 | 14.7 | 180 | CQ232723 Sequence |
| C 4 | 300 | 14.7 | 180 | CQ270754 Sequence |

| | | | | | | |
|------|------|------|-----|----|-----------|---------------------|
| C 5 | 300 | 14.7 | 180 | 6 | CQ308036 | CQ308036 Sequence |
| C 6 | 300 | 14.7 | 180 | 6 | CQ344856 | CQ344856 Sequence |
| C 7 | 300 | 10.3 | 132 | 6 | CQ058095 | CQ058095 Sequence |
| C 8 | 210 | 10.3 | 132 | 6 | CQ077416 | CQ077416 Sequence |
| C 9 | 210 | 10.3 | 132 | 6 | CQ108429 | CQ108429 Sequence |
| C 10 | 210 | 10.3 | 132 | 6 | CQ147070 | CQ147070 Sequence |
| C 11 | 210 | 10.3 | 132 | 6 | CQ206855 | CQ206855 Sequence |
| C 12 | 210 | 10.3 | 132 | 6 | CQ230294 | CQ230294 Sequence |
| C 13 | 210 | 10.3 | 132 | 6 | CQ268431 | CQ268431 Sequence |
| C 14 | 210 | 10.3 | 132 | 6 | CQ305465 | CQ305465 Sequence |
| C 15 | 85 | 4.2 | 131 | 6 | AX905781 | AX905781 Sequence |
| C 16 | 85 | 4.2 | 131 | 6 | BD041314 | BD041314 Sequence |
| C 17 | 82 | 4.0 | 173 | 6 | CQ114451 | CQ114451 Sequence |
| C 18 | 82 | 4.0 | 173 | 6 | CQ153330 | CQ153330 Sequence |
| C 19 | 82 | 4.0 | 173 | 6 | CQ236619 | CQ236619 Sequence |
| C 20 | 82 | 4.0 | 173 | 6 | CQ274216 | CQ274216 Sequence |
| C 21 | 82 | 4.0 | 173 | 6 | CQ311386 | CQ311386 Sequence |
| C 22 | 82 | 4.0 | 173 | 6 | CQ348524 | CQ348524 Sequence |
| C 23 | 79.5 | 3.9 | 169 | 6 | AR269854 | AR269854 Sequence |
| C 24 | 77.5 | 3.8 | 163 | 11 | G18760 | G18760 BMS960 cow |
| C 25 | 73 | 3.6 | 190 | 6 | AX099366 | AX099366 Sequence |
| C 26 | 73 | 3.6 | 190 | 6 | BD021982 | BD021982 Sequence |
| C 27 | 66 | 3.2 | 104 | 9 | HSEK14B01 | X92500 H. sapiens t |
| C 28 | 66 | 3.2 | 154 | 6 | AR245693 | AR245693 Sequence |
| C 29 | 66 | 3.2 | 168 | 8 | AJ719154 | AJ719154 Nicotiana |
| C 30 | 65 | 3.2 | 131 | 6 | CQ080172 | CQ080172 Sequence |
| C 31 | 65 | 3.2 | 131 | 6 | CQ113458 | CQ113458 Sequence |
| C 32 | 65 | 3.2 | 131 | 6 | CQ152338 | CQ152338 Sequence |
| C 33 | 65 | 3.2 | 131 | 6 | CQ185706 | CQ185706 Sequence |
| C 34 | 65 | 3.2 | 131 | 6 | CQ235679 | CQ235679 Sequence |
| C 35 | 65 | 3.2 | 131 | 6 | CQ273247 | CQ273247 Sequence |
| C 36 | 65 | 3.2 | 131 | 6 | CQ310696 | CQ310696 Sequence |
| C 37 | 65 | 3.2 | 131 | 6 | CQ347535 | CQ347535 Sequence |
| C 38 | 64 | 3.1 | 136 | 14 | AY366579 | AY366579 Hepatitis |
| C 39 | 63 | 3.1 | 102 | 9 | S65996 | S65996 NK-TR-putat |
| C 40 | 63 | 3.1 | 135 | 14 | AY366585 | AY366585 Hepatitis |
| C 41 | 62.5 | 3.1 | 135 | 9 | AB0046832 | AB004683 Human gen |
| C 42 | 62.5 | 3.1 | 156 | 6 | CQ397944 | CQ397944 Sequence |
| C 43 | 62.5 | 3.1 | 156 | 6 | CQ404236 | CQ404236 Sequence |
| C 44 | 62 | 3.0 | 185 | 10 | MMASPART5 | X07306 Mouse aspar |
| C 45 | 62 | 3.0 | 188 | 6 | BD071586 | BD071586 Secreted |

ALIGNMENTS

RESULT 1
CQ110706/c LOCUS CQ110706 180 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 19565 from Patent WO0157272.
ACCESSION CQ110706
VERSION CQ110706.1 GI:41080089
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human placenta
Patent: WO 0157272-A 19565 09-AUG-2001,
Acomica, Inc. (US)
FEATURES
Source location/Qualifiers
1..180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Map to AC006501.5-EXPRESSED IN PLACENTA. SIGNAL =
1.9-BEST HUMAN HIT: AL040518.1, EVALU8 1.00e-94-NT HIT:
U43139.1, EVALU8 1.10e-01-SWISSPROT HIT: Q16760, EVALU8
6.80e-02"

ORIGIN

Alignment Scores:

Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ110706 (1-180)

QY ArgArgThGluHhGlnLeuLysGlnPheThGluAspTrpSerGluGluValSer 246

Db AGCGGCAACAGACATCTGAGCAATTTACCGAATTTGTCAGAGAGATGCTCA 119

QY 247 ThTrpLeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAAAGATTTAAAGATCTTGTTGATTTTCAAGATGAATTAACATT 59

QY 267 AspGlyLysGlnLeuLeuAsnLeuThLysGluSerLeuAlaAspAspLeuLysIle 285

Db 58 GATGAAAGAAAGACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAAAATT 2

RESULT 2

CQ149443/c

LOCUS CQ149443 180 bp DNA linear PAT 21-JAN-2004

DEFINITION Sequence 19465 from Patent WO0157276.

ACCESSION CQ149443

VERSION CQ149443.1 GI:41156793

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1.180 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN BONE MARROW, SIGNAL
analysis of gene expression in human bone marrow"
Patent: WO 0157276-A 19465 09-AUG-2001;
Neomica, Inc. (US)
Location/Qualifiers
1.180

ORIGIN

Alignment Scores:

Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ149443 (1-180)

QY 227 ArgArgThGluHhGlnLeuLysGlnPheThGluAspTrpSerGluGluValSer 246

Db 178 AGCGGCAACAGACATCTGAGCAATTTACCGAATTTGTCAGAGAGATGCTCA 119

QY 247 ThTrpLeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAAAGATTTAAAGATCTTGTTGATTTTCAAGATGAATTAACATT 59

QY 267 AspGlyLysGlnLeuLeuAsnLeuThLysGluSerLeuAlaAspAspLeuLysIle 285

Db 58 GATGAAAGAAAGACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAAAATT 2

RESULT 3

CQ232723/c

CQ232723

LOCUS CQ232723 180 bp DNA linear PAT 21-JAN-2004

DEFINITION Sequence 19562 from Patent WO0157273.

ACCESSION CQ232723

VERSION CQ232723.1 GI:41215941

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

1 Penn.S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
August 2000 (03.08.00)<150> GB 24263,6<151> 03 October 2000
(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Labeling Engine
Patent: WO 0157273-A 19562 09-AUG-2001;
Neomica, Inc. (US)
Location/Qualifiers
1.180

JOURNAL

FEATURES

source

1.180 /organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN ADULT LIVER, SIGNAL
analysis of gene expression in human fetal liver"
Patent: WO 0157273-A 19015 09-AUG-2001;
Neomica, Inc. (US)
Location/Qualifiers
1.180

ORIGIN

Alignment Scores:

Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ232723 (1-180)

QY 227 ArgArgThGluHhGlnLeuLysGlnPheThGluAspTrpSerGluGluValSer 246

Db 178 AGCGGCAACAGACATCTGAGCAATTTACCGAATTTGTCAGAGAGATGCTCA 119

QY 247 ThTrpLeuCyAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAAAGATTTAAAGATCTTGTTGATTTTCAAGATGAATTAACATT 59

QY 267 AspGlyLysGlnLeuLeuAsnLeuThLysGluSerLeuAlaAspAspLeuLysIle 285

Db 58 GATGAAAGAAAGACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAAAATT 2

RESULT 4

CQ270754/c

LOCUS CQ270754 180 bp DNA linear PAT 23-JAN-2004

DEFINITION Sequence 19015 from Patent WO0157277.

ACCESSION CQ270754

VERSION CQ270754.1 GI:41243358

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

1 Penn.S.G., Hanzel,D.K., Chen,W. and Rank,D.R.
HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
ANALYSIS OF GENE EXPRESSION IN HUMAN FETAL LIVER
Patent: WO 0157277-A 19015 09-AUG-2001;
Neomica, Inc. (US)
Location/Qualifiers
1.180

FEATURES
source

Location/Qualifiers
1. 180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN FETAL LIVER, SIGNAL
= 1.3-EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"

ORIGIN

Alignment Scores:

Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ270754 (1-180)

Qy 227 ArgArgThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCySAIaGlnAAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAGATTAAAGATCTTTGGTATTTTCAAGATGAATTAACATT 59

Qy 267 AspGlyLysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuValSer 285

Db 58 GATGAGAAAGAACTGTGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAAT 2

RESULT 5

CQ308036/c
LOCUS CQ308036 180 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 19141 from Patent WO0186003.
ACCESSION CQ308036
VERSION CQ308036.1 GI:41268613

KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.

AUTHORS Human genome-derived single exon nucleic acid probes useful for

TITLE analysis of gene expression in human lung

JOURNAL Patent: WO 0186003-A 19141 15-NOV-2001;

FEATURES Location/Qualifiers

1. 180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN LUNG, SIGNAL =
0.88-EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"

ORIGIN

Alignment Scores:

Pred. No.: 6.53e-19 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ308036 (1-180)

Qy 227 ArgArgThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCySAIaGlnAAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAGATTAAAGATCTTTGGTATTTTCAAGATGAATTAACATT 59

Qy 267 AspGlyLysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuValSer 285

Db 58 GATGAGAAAGAACTGTGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAAT 2

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCySAIaGlnAAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAGATTAAAGATCTTTGGTATTTTCAAGATGAATTAACATT 59

Qy 267 AspGlyLysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuValSer 285

Db 58 GATGAGAAAGAACTGTGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAAT 2

RESULT 6

CQ344856/c
LOCUS CQ344856 180 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 18950 from Patent WO0157275.
ACCESSION CQ344856
VERSION CQ344856.1 GI:41293927

KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.

AUTHORS Human genome-derived single exon nucleic acid probes useful for

TITLE analysis of gene expression in human brain

JOURNAL Patent: WO 0157275-A 18950 09-AUG-2001;

REFERENCE 2 Aecomica, Inc. (US)

FEATURES Location/Qualifiers

source

1. 180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN BRAIN, SIGNAL =
0.73-EST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"

US-10-077-111-13 (1-384) x CQ344856 (1-180)

Qy 227 ArgArgThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValSer 246

Db 178 AGGCGCAGACAACATCGCTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119

Qy 247 ThrTrpLeuCySAIaGlnAAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIle 266

Db 118 ACATGGCTTTGTGCACAGATTAAAGATCTTTGGTATTTTCAAGATGAATTAACATT 59

Qy 267 AspGlyLysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuValSer 285

Db 58 GATGAGAAAGAACTGTGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAAT 2

RESULT 7

CQ058095/c
LOCUS CQ058095 132 bp DNA linear PAT 19-JAN-2004
DEFINITION Sequence 8915 from Patent WO0157270.
ACCESSION CQ058095
VERSION CQ058095.1 GI:41032601

KEYWORDS Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.

AUTHORS Human genome-derived single exon nucleic acid probes useful for

TITLE analysis of gene expression in human lung

JOURNAL Patent: WO 0186003-A 19141 15-NOV-2001;

FEATURES Location/Qualifiers

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human bone marrow
JOURNAL Patent: WO 0157276-A 17092 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES location/Qualifiers
source 1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN BONE MARROW. SIGNAL
= 0.75-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ALIGNMENT Scores:
Pred. No.: 8.22e-11 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ147070 (1-132)

RESULT 11
LOCUS CQ206855/c 132 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 9132 from Patent WO0157271.
ACCESSION CQ206855
VERSION CQ206855.1 GI:41192959
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human breast and DC 474 cells
JOURNAL Patent: WO 0157271-A 9132 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES location/Qualifiers
source 1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN BT474. SIGNAL =
0.7-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN
Alignment Scores:
Pred. No.: 8.22e-11 Length: 132
Score: 210.00 Matches: 43

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ206855 (1-132)

RESULT 12
LOCUS CQ230294/c 132 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 17133 from Patent WO0157273.
ACCESSION CQ230294
VERSION CQ230294.1 GI:41213512
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Listing Engine
Patent: WO 0157273-A 17133 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES location/Qualifiers
source 1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN ADULT LIVER. SIGNAL
= 0.73-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN
Alignment Scores:
Pred. No.: 8.22e-11 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CQ230294 (1-132)

RESULT 13
LOCUS CQ230294/c 132 bp DNA linear PAT 21-JAN-2004
DEFINITION Sequence 17133 from Patent WO0157273.
ACCESSION CQ230294
VERSION CQ230294.1 GI:41213512
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO
3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Listing Engine
Patent: WO 0157273-A 17133 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES location/Qualifiers
source 1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN ADULT LIVER. SIGNAL
= 0.73-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN
Alignment Scores:
Pred. No.: 8.22e-11 Length: 132
Score: 210.00 Matches: 43

DB 10 ATCGCATCA 2
RESULT 13
LOCUS CQ268431/c 132 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 16692 from Patent WO0157277.
ACCESSION CQ268431
VERSION CQ268431.1 GI:41241035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
JOURNAL analysis of gene expression in human fetal liver
Patent: WO 0157277-A 1692 09-AUG-2001;
Neomica, Inc. (US)
FEATURES
source location/Qualifiers
1..132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN FETAL LIVER, SIGNAL
= 0.78-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"
ORIGIN
Alignment Scores:
Pred. No.: 8.22e-11 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 6 Gaps: 0
US-10-077-111-13 (1-384) x CQ268431 (1-132)
QY 290 LeuArgSerLyValIeuArgLyIleGluGluLeuArgThryValLySerLeuSer 309
DB 130 CTCGCTAGTAAAGTCTGAGGAAATTGAAGAGCTCGACCAAGTTAAATCCCTTCT 71
QY 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspProVal 329
DB 70 TCAGGAATTCCTGATGAAATTATATGTCCAATTAAGAACTTATGAAGAATCCGATC 11
QY 330 IleAlaSer 332
DB 10 ATCGCATCA 2
RESULT 14
LOCUS CQ305465/c 132 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 16570 from Patent WO0186003.
ACCESSION CQ305465
VERSION CQ305465.1 GI:41266042
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE Human genome-derived single exon nucleic acid probes useful for
JOURNAL analysis of gene expression in human lung
Patent: WO 0186003-A 16570 15-NOV-2001;
Neomica, Inc. (US)
FEATURES
source location/Qualifiers
1..132
/organism="Homo sapiens"

/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN LUNG, SIGNAL =
0.76-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"
ORIGIN
Alignment Scores:
Pred. No.: 8.22e-11 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 6 Gaps: 0
US-10-077-111-13 (1-384) x CQ305465 (1-132)
QY 290 LeuArgSerLyValIeuArgLyIleGluGluLeuArgThryValLySerLeuSer 309
DB 130 CTCGCTAGTAAAGTCTGAGGAAATTGAAGAGCTCGACCAAGTTAAATCCCTTCT 71
QY 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspProVal 329
DB 70 TCAGGAATTCCTGATGAAATTATATGTCCAATTAAGAACTTATGAAGAATCCGATC 11
QY 330 IleAlaSer 332
DB 10 ATCGCATCA 2
RESULT 15
LOCUS AX905781 191 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 21644 from Patent EP1033401.
ACCESSION AX905781
VERSION AX905781.1 GI:40060762
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Dumais, Milne, Edwards, J.B., Duclet, A. and Giordano, J.Y.
TITLE Expressed sequence tags and encoded human proteins
JOURNAL Patent: EP 1033401-A 21644 06-SEP-2000;
Genet (FR)
FEATURES
source location/Qualifiers
1..191
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 42.4 Length: 191
Score: 85.00 Matches: 18
Percent Similarity: 50.00% Conservative: 7
Best Local Similarity: 36.00% Mismatches: 25
Query Match: 4.15% Indels: 0
DB: 6 Gaps: 0
US-10-077-111-13 (1-384) x AX905781 (1-191)
QY 100 ValArgValCysGlnIleHisSerProAspSerThrCysLeuAlaSerGlyValAlaAspGly 119
DB 42 GTTACTGCTGCTTCATCTCCAGACTGCAGACATCTGCTGCTGACCTGAGAGAG 101
QY 120 ThrValValLeuThrPheAlaGlnSerTyrLysLeuTyrArgCysGlySerVallyAsp 139
DB 102 TCGGTCTTCTTAAGAGCATGAGAGTCTACAGTMAATTCCGAACTAGAGGCCATCA 161
QY 140 GlySerLeuAlaAlaCysAlaPheSerPro 149

Mon Feb 7 10:23:07 2005

us-10-077-111-13.p2nsz1m200.rge

Page 7

Db 162 AGCAGTGTGTCCTTGACTTCCCCC 191

Search completed: February 5, 2005, 16:54:50
Job time : 5696 secs

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 5, 2005, 12:08:02 ; Search time 698 Seconds
(without alignments)
3256.706 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVKLIHTLADHGDVNCACF.....LTPNRITKMAINRWLETHOK 384

Scoring table: BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 segs, 2959870667 residues
Total number of hits satisfying chosen parameters: 4887668

Minimum DB seg length: 0
Maximum DB seg length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+.p2n.model -DEV=x1h
-Q=/cgr2_1/USPTO.spool/US10077111/runat_04022005_121140_7174/app_query.fasta_1.583
-DB=N_Geneseq_16Dec04 -OPMT=fastcap -SUFFIX=p2nslm200.rng -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200 -USER=US10077111_@CGN_1_1_644_@runat_04022005_121140_7174 -NCPU=6
-ICPU=3 -NO_MMAP -LARGEOUTERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| C 1 | 300 | 14.7 | 180 | 4 | ABA70710 Human foe |
| C 2 | 300 | 14.7 | 180 | 4 | AA150879 Probe #19 |
| C 3 | 300 | 14.7 | 180 | 4 | AAK44908 Human bon |
| C 4 | 300 | 14.7 | 180 | 4 | AAK18959 Human bra |
| C 5 | 300 | 14.7 | 180 | 4 | ABS44572 Human liv |

| | | | | | |
|------|------|------|-----|----|---------------------|
| C 6 | 300 | 14.7 | 180 | 6 | ABS19150 Human gen |
| C 7 | 210 | 10.3 | 132 | 4 | AA123283 Probe #13 |
| C 8 | 210 | 10.3 | 132 | 4 | ABA68387 Human foe |
| C 9 | 210 | 10.3 | 132 | 4 | AA148602 Probe #17 |
| C 10 | 210 | 10.3 | 132 | 4 | ABA50437 Human bre |
| C 11 | 210 | 10.3 | 132 | 4 | AAK42535 Human bon |
| C 12 | 210 | 10.3 | 132 | 4 | ABS42143 Human liv |
| C 13 | 210 | 10.3 | 132 | 5 | AA108924 Probe #89 |
| C 14 | 210 | 10.3 | 132 | 6 | ABS16579 Human gen |
| C 15 | 150 | 7.3 | 92 | 4 | AAH36289 Human col |
| C 16 | 93.5 | 4.6 | 174 | 12 | ACH93376 Human gen |
| C 17 | 87 | 4.3 | 150 | 12 | ACH85482 Human gen |
| C 18 | 85 | 4.2 | 191 | 3 | AACT17569 Human sec |
| C 19 | 82 | 4.0 | 173 | 4 | ABA74172 Human foe |
| C 20 | 82 | 4.0 | 173 | 4 | AA154624 Probe #23 |
| C 21 | 82 | 4.0 | 173 | 4 | AAK48795 Human bon |
| C 22 | 82 | 4.0 | 173 | 4 | AAK22627 Human bra |
| C 23 | 82 | 4.0 | 173 | 4 | ABS48468 Human liv |
| C 24 | 82 | 4.0 | 173 | 6 | ABS22500 Human gen |
| C 25 | 79.5 | 3.9 | 169 | 10 | ACA55819 C. elegans |
| C 26 | 79.5 | 3.9 | 159 | 12 | AD155615 Human pol |
| C 27 | 78.5 | 3.8 | 155 | 12 | ACH82916 Human gen |
| C 28 | 73 | 3.6 | 190 | 2 | AAV11599 Homo sapi |
| C 29 | 73 | 3.6 | 190 | 2 | AAV11599 Homo sapi |
| C 30 | 72.5 | 3.5 | 197 | 6 | AAV98377 Human CDN |
| C 31 | 72.5 | 3.5 | 197 | 6 | ABV89016 Human col |
| C 32 | 67.5 | 3.5 | 197 | 6 | ABV66894 Human col |
| C 33 | 67.5 | 3.5 | 41 | 4 | AA164919 Beta-cran |
| C 34 | 66 | 3.2 | 131 | 12 | ACH92306 Human gen |
| C 35 | 65.5 | 3.2 | 154 | 10 | ABX82592 Corn ear- |
| C 36 | 65 | 3.2 | 167 | 6 | ABV96402 Human pan |
| C 37 | 65 | 3.2 | 131 | 4 | AA126039 Probe #15 |
| C 38 | 65 | 3.2 | 131 | 4 | ABA73203 Human foe |
| C 39 | 65 | 3.2 | 131 | 4 | AA153631 Probe #22 |
| C 40 | 65 | 3.2 | 131 | 4 | ABA48636 Probe #17 |
| C 41 | 65 | 3.2 | 131 | 4 | AAK47803 Human bon |
| C 42 | 65 | 3.2 | 111 | 4 | AAK21638 Human bra |
| C 43 | 65 | 3.2 | 111 | 6 | ABS47528 Human liv |
| C 44 | 64 | 3.1 | 118 | 12 | ABS21810 Human gen |
| C 45 | 64 | 3.1 | 142 | 12 | ACH83318 Human gen |

ALIGNMENTS

RESULT 1
ID ABA70710 standard; DNA; 180 BP.
XX
AC ABA70710;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #19015.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLB-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-483447/52.
XX
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
PS Claim 4; SEQ ID NO 19015; 6339p + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at fcp.wipo.int/pub/published_pcr_sequences
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: Gaps: 0
US-10-077-111-13 (1-384) x ABA70710 (1-180)
QY 227 ArgArgThrGluHieGlnLeuLysGlnPheThrGluAptTrpSerGluGluValSer 246
Db 178 AGGCGCACAGAACATCAGTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119
QY 247 ThTTrpLeuCYaAlaGlnAapLeuLysAapLeuValGlyIlePheLysMetAenAenIle 266
Db 118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGTGGTATTTTCAAGATGAATACATT 59
QY 267 AapGlyLYeGlnLeuLeuAenLeuThrLYeGlnSerLeuAlaAapAapLeuLysIle 285
Db 58 GATGAAAAGAACTGTTGAATCTTACAAAGAAAGTGGCTGATGATTTGAAAATT 2
RESULT 2
AA150879/c
ID AA150879 standard; DNA; 180 BP.
XX
AC AA150879;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #19565 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN W0200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
XX
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 19565; 6549p; English.
XX
CC The present invention relates to single exon nucleic acid probes (ssNP).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful
CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: Gaps: 0
US-10-077-111-13 (1-384) x AA150879 (1-180)
QY 227 ArgArgThrGluHieGlnLeuLysGlnPheThrGluAptTrpSerGluGluValSer 246
Db 178 AGGCGCACAGAACATCAGTGAAGCAATTACCGAAGATTGTCAGAGAGATGCTCA 119
QY 247 ThTTrpLeuCYaAlaGlnAapLeuLysAapLeuValGlyIlePheLysMetAenAenIle 266
Db 118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGTGGTATTTTCAAGATGAATACATT 59
QY 267 AapGlyLYeGlnLeuLeuAenLeuThrLYeGlnSerLeuAlaAapAapLeuLysIle 285
Db 58 GATGAAAAGAACTGTTGAATCTTACAAAGAAAGTGGCTGATGATTTGAAAATT 2
RESULT 3
AAK44908/c
ID AAK44908 standard; DNA; 180 BP.
XX
AC AAK44908;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 19465.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN W0200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX

DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 19465; 658bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 4 Gaps: 0
US-10-077-111-13 (1-384) x AAK44908 (1-180)
QY 227 ArgArGThrGluHISGlnLeuLySGlnPheThrGluAspTrpSerGluValValSer 246
DB 178 AGGGGCAACAGAACATCAGCTGAGCAATTACCGAAGATTGGTCAGAGAGATGCTCA 119
QY 247 ThrTrpLeuCySAIaGlnAspLeuLyAspLeuValGlyIlePheLyMetAsnAsnIle 266
DB 118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGCTATTTCACAGATGAATTAACATT 59
QY 267 AspGlyLySGlnLeuLeuAsnLeuThrLySGlnSerLeuAlaAspAspLeuLyIle 285
DB 58 GATGGAAGAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTGAAATTT 2
RESULT 4
AAK18959/c
ID AAK18959 standard; DNA; 180 BP.
XX
AC AAK18959;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe SEQ ID NO: 18950.
XX
KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
KW ss.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000667.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.

XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 18950; 650bp + Sequence listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is one of the probes of the
CC invention
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: 4 Gaps: 0
US-10-077-111-13 (1-384) x AAK18959 (1-180)
QY 227 ArgArGThrGluHISGlnLeuLySGlnPheThrGluAspTrpSerGluValValSer 246
DB 178 AGGGGCAACAGAACATCAGCTGAGCAATTACCGAAGATTGGTCAGAGAGATGCTCA 119
QY 247 ThrTrpLeuCySAIaGlnAspLeuLyAspLeuValGlyIlePheLyMetAsnAsnIle 266
DB 118 ACATGGCTTTGTGCACAAAGATTAAAGATCTTGCTATTTCACAGATGAATTAACATT 59
QY 267 AspGlyLySGlnLeuLeuAsnLeuThrLySGlnSerLeuAlaAspAspLeuLyIle 285
DB 58 GATGGAAGAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTGAAATTT 2
RESULT 5
ABS44572/c
ID ABS44572 standard; DNA; 180 BP.
XX
AC ABS44572;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID NO 19562.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
XX Claim 4; SEQ ID NO 19562; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/fragments). The probe hybridizes at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: Gaps: 0
US-10-077-111-13 (1-384) x ABS44572 (1-180)
QY ArgArGThGluHISGlnLeuLYSGlnPheThGluAAPTTPSerGluGluValSer 246
DB 178 AGCGGCACAGAACATCGCTGAAACATTACCGAAGATTGCTCAGGAGATGTCTCA 119
QY 247 ThTTPLeuCYsaIaGlnAaPLeuLYsAaPLeuValGlyIlePheLYsMeLaMaAnIle 266
DB 118 ACATGGCTTGTGCACAGATTAAAGATCTGTGATTATTTTCAAGATGAATTAACATT 59
QY 267 AaGGLyLYsGlnLeuAaAnLeuThLYsGlnSerLeuAaAaPLeuLYsIle 285
DB 58 GATGGAAGAAAGACTGTGAATCTTAAGAAAGAGTGTGCTGATGATTTGAAAAATT 2
RESULT 6
ABS19150/C
ID ABS19150 standard; DNA; 180 BP.
XX
AC ABS19150;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 19141.
XX
XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hereditary-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
PN
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX

PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR,
XX WPI, 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples.
XX
XX Claim 4; SEQ ID NO 19141; 634bp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridize at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray, assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hereditary-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Kartagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1.15e-20 Length: 180
Score: 300.00 Matches: 58
Percent Similarity: 98.31% Conservative: 0
Best Local Similarity: 98.31% Mismatches: 1
Query Match: 14.66% Indels: 0
DB: Gaps: 0
US-10-077-111-13 (1-384) x ABS19150 (1-180)
QY 227 ArgArGThGluHISGlnLeuLYSGlnPheThGluAAPTTPSerGluGluValSer 246
DB 178 AGCGGCACAGAACATCGCTGAAACATTACCGAAGATTGCTCAGGAGATGTCTCA 119

| | | | |
|------------------------|---|---|-----|
| Qy | 247 | ThcTrrpleucYvAlIagIcAspleuYsAspleuVAlGyIilepelysMeAsAsnIle | 266 |
| Db | 118 | ACATGGCTTTGTGCACAAAGATTAAAAAGATCTTGTTGGTATTTCACAGATGAATDCAATT | 59 |
| Qy | 267 | AspGcLYsgsIuDeuIeuanIeuthrLYsgIuSerIeuaIAspAspleuLYsIle | 285 |
| Db | 58 | GATGCAAAAGCACTGTTGATCTTACAAAGAAAGACTGCGCTGATGATTTGAAATT | 2 |
| RESULT 7 | | | |
| ID | AA123283/c | AA123283 standard; DNA; 132 BP. | |
| AC | AA123283; | | |
| XX | 12-OCT-2001 | (first entry) | |
| DT | | | |
| XX | | | |
| DE | Probe #13216 for gene expression analysis in human cervical cell sample. | | |
| XX | | | |
| KW | Probe; human; microarray; gene expression; cervical epithelial cell; | | |
| KW | cervical cancer; ss. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| PN | W0200157278-A2. | | |
| XX | | | |
| PD | 09-AUG-2001. | | |
| XX | | | |
| PF | 30-JAN-2001; 2001MO-US000670. | | |
| XX | | | |
| PR | 04-FEB-2000; 2000US-0180312P. | | |
| PR | 26-MAY-2000; 2000US-0207456P. | | |
| PR | 30-JUN-2000; 2000US-00608408. | | |
| PR | 03-AUG-2000; 2000US-00632366. | | |
| PR | 21-SEP-2000; 2000US-0234687P. | | |
| PR | 27-SEP-2000; 2000US-0236359P. | | |
| PR | 04-OCT-2000; 2000GB-00024263. | | |
| XX | | | |
| PA | (MOLE-) MOLECULAR DYNAMICS INC. | | |
| XX | | | |
| PI | Penn SG, Hanzel DK, Chen W, Rank DR; | | |
| XX | | | |
| DR | WPI; 2001-488901/53. | | |
| XX | | | |
| PT | Human genome-derived single exon nucleic acid probes useful for analyzing | | |
| PT | gene expression in human cervical epithelial cells. | | |
| XX | | | |
| PS | Claim 25; SEQ ID NO 13216; 487bp; English. | | |
| XX | | | |
| CC | The present invention relates to human single exon nucleic acid probes | | |
| CC | (SENP). The present sequence is one such probe. The SENPs are derived | | |
| CC | from human HeLa cells. The SENPs can be used to produce a single exon | | |
| CC | microarray, which can be used for measuring human gene expression in a | | |
| CC | sample derived from human cervical epithelial cells. By measuring gene | | |
| CC | expression, the probes are therefore useful in grading and/or staging of | | |
| CC | diseases of the cervix, notably cervical cancer. Note: The sequence data | | |
| CC | for this patent did not form part of the printed specification, but was | | |
| CC | obtained in electronic format directly from WIPO at | | |
| CC | ftp.wipo.int/pub/published_pct_sequences | | |
| XX | | | |
| SQ | Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other; | | |
| Alignment Scores: | | | |
| Pred. No.: | 5,44e-12 | Length: 132 | |
| Score: | 210.00 | Matches: 43 | |
| Percent Similarity: | 100.00% | Conservative: 0 | |
| Best Local Similarity: | 100.00% | Mismatches: 0 | |
| Query Match: | 10.26% | Indels: 0 | |
| DB: | 4 | Gaps: 0 | |
| Qy | US-10-077-111-13 (1-384) x AA123283 (1-132) | | |
| Qy | 290 LeuArGserLYsValLeuArGlySIleGluGluLeuArGmThrLYsValLYsSerLeuSer | | 309 |

```

Db      130 CTCGGTAGTAAGTCTGTGAGAAATTGAAAGCTCAGAGCAACGTTAAATCCCTTCT 71
Qy      310 SerGlyIleProAspGluPheIleCysProIleThrArgGluMetLysAspProVal 329
Db      70 TCAGGAATTCGATGATATTATATGTCCAATTAAGAACTTATGAAAGATCCGATC 11
Qy      330 ILeAlaSer 332
Db      10 ATCGCATCA 2

RESULT 8
ABA68387/c
ID      ABA68387 standard; DNA; 132 BP.
XX
AC      ABA68387;
XX
DT      01-FEB-2002 (first entry)
XX
DE      Human foetal liver single exon nucleic acid probe #16692.
XX
KW      Human; foetal liver; gene expression; single exon nucleic acid probe; ss
XX
OS      Homo sapiens.
XX
WO200157277-A2.
XX
PD      09-AUG-2001.
XX
PF      30-JAN-2001; 2001WO-US000669.
XX
PR      04-FEB-2000; 2000US-0180312P.
PR      26-MAY-2000; 2000US-0207456P.
PR      30-JUN-2000; 2000US-00608408.
PR      03-AUG-2000; 2000US-00632366.
PR      21-SEP-2000; 2000US-0234687P.
PR      27-SEP-2000; 2000US-0236359P.
PR      04-OCT-2000; 2000GB-00024263.
XX
PA      (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI      Penn SG, Hanzel DK, Chen W, Rank DR;
DR      WPI; 2001-483447/52.
XX
PT      Human genome-derived single exon nucleic acid probes useful for analyzing
PT      gene expression in human foetal liver.
XX
PS      Claim 4; SEQ ID NO 16692; 639pp + Sequence Listing; English.
XX
CC      The invention relates to a single exon nucleic acid probe for measuring
CC      human gene expression in a sample derived from human foetal liver. The
CC      single exon nucleic acid probes may be used for predicting, measuring and/or
CC      displaying gene expression in samples derived from human foetal liver. The
CC      present sequence is a single exon nucleic acid probe of the invention.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
CC      at ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      5.44e-12      Length:      132
Score:          210.00      Matches:      43
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      10.26%      Indels:      0
DB:              Gaps:      0

US-10-077-111-13 (1-384) x ABA68387 (1-132)
Qy      290 LeuArgSerIlyeValLeuArgIlyValIleGluGluLeuArgThrIlyeValIlyeSerIleuSer 309

```

Db 130 CTCGGTAGTAAGTCTGAGGAAATTGAAAGACTCAGACCAAGTTAAATCCCTTCT 71
Qy 310 SerGly11eProAspGluPheIleCySProlleThrArgGluLeuMetLysAspProval 329
Db 70 TCAGGAATTCCTGATGATTATATGTCCAATTAAGTAACTTATGAAAGATCCGGTC 11
Qy 330 IleAlaSer 332
Db 10 ATCGCATCA 2
RESULT 9
AA148602/C
ID AA148602 standard; DNA; 132 BP.
AC AA148602;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #17288 used to measure gene expression in human placenta sample.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
XX Homo sapiens.
XX
XX WO200157272-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-488937/53.
XX DR
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX
XX PS Claim 25; SEQ ID NO 17288; 654bp; English.
XX
XX CC The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.44e-12 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 4 Gaps: 0
US-10-077-111-13 (1-384) x AA148602 (1-132)
Qy 290 LeuArgSerIysValIleuArgLysIleGluGluLeuArgThrIysValIysSerLeuSer 309
Db 130 CTCGGTAGTAAGTCTGAGGAAATTGAAAGACTCAGACCAAGTTAAATCCCTTCT 71
Qy 310 SerGly11eProAspGluPheIleCySProlleThrArgGluLeuMetLysAspProval 329

Db 70 TCAGGAATTCCTGATGATTATATGTCCAATTAAGTAACTTATGAAAGATCCGGTC 11
Qy 330 IleAlaSer 332
Db 10 ATCGCATCA 2
RESULT 10
ABA50437/C
ID ABA50437 standard; DNA; 132 BP.
XX
XX ABA50437;
XX
XX 01-FEB-2002 (first entry)
XX
XX Human breast cell single exon nucleic acid probe #9132.
XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200157271-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US000662.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-496933/54.
XX DR
XX PT New spatially-addressable set of single exon nucleic acid probes, useful
XX PT for measuring gene expression in sample derived from human breast,
XX PT comprises number of single exon nucleic acid probes.
XX
XX PS Claim 4; SEQ ID NO 9132; 327bp + Sequence Listing; English.
XX
XX CC The invention relates to a spatially-addressable set of single exon
XX CC nucleic acid probes for measuring gene expression in a sample derived
XX CC from human breast and BT 474 cells. The method involves contacting the
XX CC probes with a collection of detectably labelled nucleic acids derived
XX CC from mRNA of human breast, and then measuring the label bound to each
XX CC probe of the microarray. The probes are useful for verifying the
XX CC expression of regions of genomic DNA predicted to encode proteins. They
XX CC are useful for gene discovery, and for determining predisposition and/or
XX CC prognosing breast disease. Gene expression analysis is useful for
XX CC assessing the toxicity of chemical agents on cells. The microarray of
XX CC this invention presents a far greater diversity of probes for measuring
XX CC gene expression, with far less bias than expressed sequence tag
XX CC microarrays. The method is suitable for rapid production of functional
XX CC information from genomic sequence. The present sequence is a single exon
XX CC nucleic acid probe of the invention. Note: The sequence data for this
XX CC patent did not form part of the printed specification, but was obtained
XX CC in electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 5.44e-12 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0
DB: 4 Gaps: 0

US-10-077-111-13 (1-384) x ABA50437 (1-132)

QY 290 LeuArgSerIysValLeuArgIysIleGIuGIuLeuArgThrIysValIysSerLeuSer 309
DB 130 CTGGCTAGTAAAGCTGAGGAAATTGAAGAGCTCAGACCAAGTTAAATCCCTTCT 71

QY 310 SerGIyIleProAspGluPheIleCysProIleThrArgGIuLeuMetIysAspProVal 329
DB 70 TCAGGAATTCCTGATGAAATTATATGTCCAATAAAGAGAACTTATGAAAGATCCGCTC 11

QY 330 IleAlaSer 332
DB 10 ATCGCATCA 2

RESULT 11
ID AAK42535/c
ID AAK42535 standard; DNA; 132 BP.

AC AAK42535;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 17092.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 17092; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.44e-12 Length: 132
Score: 210.00 Matches: 43
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.26% Indels: 0

DB: 4 Gaps: 0

US-10-077-111-13 (1-384) x AAK42535 (1-132)

QY 290 LeuArgSerIysValLeuArgIysIleGIuGIuLeuArgThrIysValIysSerLeuSer 309
DB 130 CTGGCTAGTAAAGCTGAGGAAATTGAAGAGCTCAGACCAAGTTAAATCCCTTCT 71

QY 310 SerGIyIleProAspGluPheIleCysProIleThrArgGIuLeuMetIysAspProVal 329
DB 70 TCAGGAATTCCTGATGAAATTATATGTCCAATAAAGAGAACTTATGAAAGATCCGCTC 11

QY 330 IleAlaSer 332
DB 10 ATCGCATCA 2

RESULT 12
ID ABS42143/c
ID ABS42143 standard; DNA; 132 BP.

AC ABS42143;
XX
DT 25-FEB-2003 (first entry)
XX
DE Human liver single exon probe, SEQ ID No 17133.
XX
KW Human; single exon nucleic acid probe; liver; cirrhosis;
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KW coronary heart disease; ss.
XX
OS Homo sapiens.
XX
PN WO200157273-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000664.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-48898/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 17133; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5,44e-12 | Length: | 132 |
| Score: | 210.00 | Matches: | 43 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 10.26% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-10-077-111-13 (1-384) x ABS42143 (1-132)

QY 290 LeuArgSerIysValIleuAglYsIIeGIuGluLeuArgThrIyValIysSerLeuSer 309
|||
DB 130 CTCGGTAGTAAAGTGTGAGGAAATTTGAAGAGCTCAGGACCAAGGTTAAATCCCTTCT 71
|||
QY 310 SerGIYIleProApgIuPheIleCySProlleThArGIuLeuWeIySApProVal 329
|||
DB 70 TCGAGAAATTCCTGATGATTTATATGTCCAATTAAGAACTTATGAAAGATCCGGTC 11
|||
QY 330 IleAlaSer 332
|||
DB 10 ATCGCATCA 2

RESULT 13
AA108924/C
ID AA108924 standard; DNA; 132 BP.

AC AA108924;
XX
DT 09-OCT-2001 (first entry)
XX
DE Probe #8915 used to measure gene expression in human breast sample.
XX
KM Probe: human; breast disease; breast cancer; development disorder; ss;
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
OS Homo sapiens.
XX
PN WO200157270-A2.
XX
PD 09-AUG-2001.
XX
PF 29-JAN-2001; 2001WO-US000661.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
PT Novel single exon nucleic acid probe used to measuring gene expression in
XX a human breast.
XX
PS Claim 25; SEQ ID NO 8915; 322dp; English.

CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form

CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5,44e-12 | Length: | 132 |
| Score: | 210.00 | Matches: | 43 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 10.26% | Indels: | 0 |
| DB: | 5 | Gaps: | 0 |

US-10-077-111-13 (1-384) x AA108924 (1-132)

QY 290 LeuArgSerIysValIleuAglYsIIeGIuGluLeuArgThrIyValIysSerLeuSer 309
|||
DB 130 CTCGGTAGTAAAGTGTGAGGAAATTTGAAGAGCTCAGGACCAAGGTTAAATCCCTTCT 71
|||
QY 310 SerGIYIleProApgIuPheIleCySProlleThArGIuLeuWeIySApProVal 329
|||
DB 70 TCGAGAAATTCCTGATGATTTATATGTCCAATTAAGAACTTATGAAAGATCCGGTC 11
|||
QY 330 IleAlaSer 332
|||
DB 10 ATCGCATCA 2

RESULT 14
ABS16579/C
ID ABS16579 standard; DNA; 132 BP.

AC ABS16579;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 16570.
XX
XX
KM Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
KM chronic obstructive pulmonary disease; interstitial lung disease;
KM familial idiopathic pulmonary fibrosis; neurofibromatosis;
KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KM Hereditary-Pudlak syndrome; sarcoidosis; pulmonary haemostelerosis;
KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karsenger syndrome;
KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KM primary ciliary dyskinesia; pulmonary hypertension;
KM hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 16570; 634dp; English.

XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridize at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridization of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridization to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemangioma, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsagen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5.44e-12 | Length: | 132 |
| Score: | 210.00 | Matches: | 43 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 10.26% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-077-111-13 (1-384) x ABS16579 (1-132)

QY 290 LeuArgSerIyValLeuArgIyIeGlulGluLeuArgThlyValIySerIeuSer 309

DB 130 CTGCGTGTAAAGCGTCAGAGAAATTCAGAGCTCAAGCAAGTTAAATCCCTTTCT 71

QY 310 SerGlyIleProApgIuPhelIeCyseProIleThArgGluLeuMetIyAspProVal 329

DB 70 TCAGGAATTCCTGATGAAATTAATGTCACATACTAGAGAACTTAAGAAATCCGCTC 11

QY 330 IleaIaSer 332

DB 10 ATCGCATCA 2

RESULT 15
ID AAH36289 standard; cDNA; 92 BP.
AC AAH36289;
XX
DT 03-SEP-2001 (first entry)
XX
DE Human colon cancer antigen encoding cDNA SEQ ID NO:3371.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; ss.
XX Homo sapiens.
OS
PN W0200122920-A2.
PM
XX
PD 05-APR-2001.
PF 28-SEP-2000; 2000MO-US026524.
PR 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX
DR WPI: 2001-235357/24.
DR P-PSDB; AAG76884.
PT Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.
PS Claim 1; Page 5177-5178; 9803pp; English.

XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB7789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922

XX SQ Sequence 92 BP; 27 A; 9 C; 23 G; 28 T; 0 U; 5 Other;

Alignment Scores:

| | | | |
|------------------------|----------|---------------|----|
| Pred. No.: | 2.66e-06 | Length: | 92 |
| Score: | 150.00 | Matches: | 26 |
| Percent Similarity: | 92.86% | Conservative: | 0 |
| Best Local Similarity: | 92.86% | Mismatches: | 2 |
| Query Match: | 7.33% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-10-077-111-13 (1-384) x AAH36289 (1-92)

QY 161 GlYAspLeuThrValITTPAsPAsPlyMeLArgCyseIuHIsSerGluIyValIaHIsAsp 180

DB 1 GGTGATTTAAACAGTGTGCGATGATTAATAAGTGTGTGATGNGAAGCAACATGAT 60

QY 181 LeuGlyIleThrCyseAspPhe 188

DB 61 CTTGGAATTAACCTGCTGNGATTTT 84

Search completed: February 5, 2005, 15:19:50
Job time : 701 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: February 5, 2005, 14:52:43 ; Search time 4204 Seconds
(without alignments)
3476.850 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047

Sequence: 1 MVXLIHTLMDHGDVNCACF.....LTPNRLTKAIAINRLETHOK 384

Scoring table:

BLISSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 3966416

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US10077111/runat_04022005_121141_7194/app_query.fasta.1.583
-DB=EST -OPMT=fastcap -SUFFIX=pnszlm200.rst -MINMATCH=0.1 -LOOPCT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blissum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200
-USRR=US10077111.qcgn_1_1_4352.qrunat_04022005_121141_7194 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEROVERLY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: EST:*

1: gb_est1.*
2: gb_est2.*
3: gb_hlc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g981.*
9: gb_g982.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|-------|-------------|
| C 1 | 345 | 16.9 | 199 2 | BE669639 |
| C 2 | 342 | 16.7 | 199 2 | BF110272 |
| C 3 | 290 | 14.2 | 160 1 | AA501581 |
| C 4 | 289 | 14.1 | 199 1 | AA1415375 |
| C 5 | 209.5 | 10.2 | 190 2 | BF653223 |
| C 6 | 174 | 8.5 | 138 6 | CA587431 |
| C 7 | 151 | 7.4 | 191 2 | BB323974 |
| C 8 | 146.5 | 7.2 | 144 1 | AI622018 |
| C 9 | 138 | 6.7 | 179 9 | CC724603 |

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|-------|-------------|
| C 10 | 136 | 6.6 | 82 1 | AA912881 |
| C 11 | 129 | 6.3 | 184 9 | AG211342 |
| C 12 | 122 | 6.0 | 175 2 | BE240075 |
| C 13 | 117 | 5.7 | 145 9 | CG460443 |
| C 14 | 113 | 5.5 | 189 2 | AM032239 |
| C 15 | 111 | 5.4 | 177 2 | BB186455 |
| C 16 | 108 | 5.3 | 131 9 | CG460401 |
| C 17 | 103.5 | 5.1 | 195 1 | AI202087 |
| C 18 | 99 | 4.8 | 153 5 | BO760356 |
| C 19 | 95.5 | 4.7 | 102 7 | CR455364 |
| C 20 | 88 | 4.3 | 182 7 | CO261506 |
| C 21 | 87.5 | 4.3 | 174 7 | CNS79930 |
| C 22 | 86.5 | 4.2 | 181 2 | BE461636 |
| C 23 | 86 | 4.2 | 132 1 | AF071906 |
| C 24 | 85 | 4.2 | 172 1 | AF071907 |
| C 25 | 85 | 4.2 | 139 7 | CK617095 |
| C 26 | 84.5 | 4.1 | 199 5 | BW563342 |
| C 27 | 84 | 4.1 | 131 2 | BE320194 |
| C 28 | 83 | 4.1 | 198 7 | CO310903 |
| C 29 | 82.5 | 4.0 | 197 9 | CR399937 |
| C 30 | 81 | 4.0 | 198 1 | AV902435 |
| C 31 | 80 | 3.9 | 187 6 | CD135364 |
| C 32 | 79.5 | 3.9 | 168 5 | BQ569612 |
| C 33 | 79 | 3.9 | 183 4 | BG382358 |
| C 34 | 78.5 | 3.8 | 196 8 | BH197534 |
| C 35 | 78.5 | 3.8 | 197 2 | BF472123 |
| C 36 | 78 | 3.8 | 168 4 | BG009910 |
| C 37 | 78 | 3.8 | 160 7 | CO427973 |
| C 38 | 77.5 | 3.8 | 169 7 | T82938 |
| C 39 | 76.5 | 3.7 | 178 7 | CV297531 |
| C 40 | 76 | 3.7 | 172 1 | AI369463 |
| C 41 | 76 | 3.7 | 187 8 | BH853006 |
| C 42 | 76 | 3.7 | 189 7 | N55666 |
| C 43 | 75.5 | 3.7 | 173 1 | AL785118 |
| C 44 | 75 | 3.7 | 163 7 | W17747 |
| C 45 | 75 | 3.7 | 198 1 | AA049560 |

ALIGNMENTS

RESULT 1
BE669639/c
LOCUS
DEFINITION
BE669639 199 bp. mRNA. linear. EST 08-SEP-2000
7e16e07.x1 NCI CGAP lu24 Homo sapiens CDNA clone IMAGE:3282660 3'
Similar to TR:080828 080828 HYPOTHETICAL 88.8 KD PROTEIN. ;, mRNA
Sequence.
ACCESSION
BE669639
VERSION
BE669639.1 GI:10030180
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 199)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.
Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.
Email: cgaaps-remail.nih.gov
Contact: Robert Strausberg, Ph.D.
Unpublished (1997)
Tumor Gene Index
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
CONTACT: Robert Strausberg, Ph.D.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LIML, send email to:
info@image.liml.gov
Trace considered overall poor quality
Seq primer: -400P from Gibco
High quality sequence stop: 1.
Location/Qualifiers
1. 199

FEATURES
source

```
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="IMAGE:3282660"  
/issue_type="carcinoid"  
/lab_host="DH108"  
/clone_1lb="NCI_CGAP Lu24"  
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a  
modified polylinker; Plasmid DNA from the normalized  
library NCI CGAP Lu5 was prepared, and ss circles were  
made in vitro. Following HAP purification, this DNA was  
used as tracer in a subtractive hybridization reaction.  
The driver was PCR-amplified cDNAs from a pool of 5,000  
clones made from the same library (cloneids  
141920-1417991 and 1520904-1522439). Subtraction by Bento  
Soares and M. Fatima Bonaldo."
```

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 2,74e-25 | Length: | 199 |
| Score: | 345.00 | Matches: | 66 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 16.85% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-10-077-111-13 (1-384) x BE669639 (1-199)

QY 316 PheIIcysProlleThrArgGluLeuMetLysAspProValIleAlaSerAspGlyThr 335

DB 199 TTTATATGCTCAATTAAGTAAGAACTTATGAAATCGGATCGATCGATCGAATGGTAT 140

QY 336 SerTyrGluLysGluAlaMetGluAsnTrpIleSerLysLysArgThrSerProMet 355

DB 139 TCATATGAAAAGAGAGCAATGAAATTGATCAGCAAAAAGAACTACAAGTCCATG 80

QY 356 ThrAsnLeuValLeuProSerAlaValLeuThrProAsnArgThrLeuLysMetAlaIle 375

DB 79 ACAAACTTGTCTCTCTCTCAGCGGTACTTACACCAATAGGAACTGAAAATGGCCATC 20

QY 376 AsnArgTrpLeuGluThr 381

DB 19 AATAGATGCTGGAGACA 2

RESULT 2
BF110272/c 199 bp mRNA linear EST 20-OCT-2000
LOCUS 7n51g07.x1 NCI CGAP Lu24 Homo sapiens cDNA clone IMAGE:3568309 3,
DEFINITION similar to TR:Q95VC6 Q95VC6 HYPOTHETICAL 48.9 KD PROTEIN. ;, mRNA
sequence.
BF110272
EST. BF110272.1 GI:10939962

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 199)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL, send email to:
info@image.lnl.gov
Trace considered overall poor quality

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Seq primer: -40UP from Gibco
High quality sequence setop: 1.
Location/Qualifiers
1. 199
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3568309"
/issue_type="carcinoid"
/lab_host="DH108"
/clone_1lb="NCI_CGAP Lu24"
/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Plasmid DNA from the normalized
library NCI CGAP Lu5 was prepared, and ss circles were
made in vitro. Following HAP purification, this DNA was
used as tracer in a subtractive hybridization reaction.
The driver was PCR-amplified cDNAs from a pool of 5,000
clones made from the same library (cloneids
141920-1417991 and 1520904-1522439). Subtraction by Bento
Soares and M. Fatima Bonaldo."

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5,53e-25 | Length: | 199 |
| Score: | 342.00 | Matches: | 66 |
| Percent Similarity: | 100.00% | Conservative: | 0 |
| Best Local Similarity: | 100.00% | Mismatches: | 0 |
| Query Match: | 16.71% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-10-077-111-13 (1-384) x BF110272 (1-199)

QY 312 IleProAspGluPheIIcysProlleThrArgGluLeuMetLysAspProValIleAla 331

DB 199 ATTCCGATGGAATTTATATGCTCAATTAAGTAAGAACTTATGAAATCGGATCGCA 140

QY 332 SerAspGlyTyrSerTyrGluLysGluAlaMetGluAsnTrpIleSerLysLysArg 351

DB 139 TCAGATGGCTATTCATATGAAAAGAACTTATGATCGCAAAAAGAAACGT 80

QY 352 ThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsnArgThrLeu 371

DB 79 ACAAGTCCATGCAATCTGTCTCTCAGCGGTACTTACACCAATAGGAACTGACTCTG 20

QY 372 LysMetAlaIleAsnArg 377
DB 19 AAAATGGCCATCAATAGA 2

RESULT 3
AA501581/c 160 bp mRNA linear EST 19-AUG-1997
LOCUS ne97g11.g1 NCI CGAP Kid1 Homo sapiens cDNA clone IMAGE:912260
DEFINITION similar to SW:FKM1 THECU P49695 PUTATIVE SERINE/THREONINE-PROTEIN
KINASE FKMA ;, mRNA sequence.
AA501581
EST. AA501581.1 GI:2236548

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 160)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs-remail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: David B. Kitzman, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at: www.bio.llnl.gov/bdrrp/image/image.html

Trace considered overall poor quality
Insert Length: 519 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

```
1. 160
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:912260"
/issue_type="kidney"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid1"
/note="Vector: pAMP10; mRNA made from invasive kidney tumor, cDNA made by oligo-dT priming. Non-directionally cloned. Size-selected on agarose gel, average insert size 600 bp. Reference: Krizman et al. (1996) Cancer Research 56:5380-5383."
```

ORIGIN

Alignment Scores:

| Pred. No.: | 7,91e-20 | Length: | 160 |
|------------------------|----------|---------------|-----|
| Score: | 290.00 | Matches: | 51 |
| Percent Similarity: | 96.23% | Conservative: | 0 |
| Best Local Similarity: | 96.23% | Mismatches: | 2 |
| Query Match: | 14.17% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-077-111-13 (1-384) x AA501581 (1-160)

```
QY 47 HisSerProLeuLysPheHisThrTyrAlaValAlaHisCysCysPheSerProSerGly 66
DB 160 CATCTTCATGGAAGTTTCACTACCTATGCTGCCACATGCTGTTATATCCCTTCAGGA 101
QY 67 HisIleLeuAlaSerCysSerThrAspGlyThrThrValLeuThrPantThrGluAlaGly 86
DB 100 CATATTTTGCATCGTGTTCACAGATGATGACCACTGTCTATGGAATACTGAAATGGA 41
QY 87 GluMetLeuAlaValMetGluGlnProSerGlySerPro 99
DB 40 CAGATGCTGCAGATGGAACAGCCCTAGTGCAGCCCT 2
```

RESULT 4

AI415375/c

199 bp mRNA linear EST 09-FEB-1999
mc35f10.x1 Soares mouse p3NMF19.5 Mus musculus cDNA clone
IMAGE:350539.3 similar to TR:023121 023121 F10G19.3 PROTEIN. ;

ACCESSION

AI415375

AI415375.1 GI:4258879

EST.

Mus musculus (house mouse)

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Contact: Maria M/Mashu-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouesest@watson.wustl.edu

This clone is available royalty-free through LINL; contact the IMAGE Consortium (info@image.llnl.gov) for further information. This clone was previously sequenced on the 5' end only, this new data is from the 3' end

Trace considered overall poor quality
Possible reversed clone: similarity on wrong strand
High quality sequence stop: 1.
Location/Qualifiers

FEATURES

source

```
1. 199
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:350539"
/dev_stage="19.5 dpc total fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares mouse p3NMF19.5"
/note="Vector: p773D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer (5' TGTTCACCAATCGTAGTGGAGCGCGCATTTTTTTTTTTT 3'), double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified p773 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. RNA was kindly provided by Dr. Minoru Ko (Wayne State University)."
```

ORIGIN

Alignment Scores:

| Pred. No.: | 1.41e-19 | Length: | 199 |
|------------------------|----------|---------------|-----|
| Score: | 289.00 | Matches: | 55 |
| Percent Similarity: | 92.42% | Conservative: | 6 |
| Best Local Similarity: | 83.33% | Mismatches: | 5 |
| Query Match: | 14.12% | Indels: | 0 |
| DB: | 1 | Gaps: | 0 |

US-10-077-111-13 (1-384) x AI415375 (1-199)

```
QY 312 IleProAspGluPheIleCysProlIleThrAspGluLeuMetLysAspProValIleAla 331
DB 199 ATCCCTGACGACGTTCAATGTCACAAATACAGAGAACTCATGAGAACCCCGTCATCGCA 140
QY 332 SerAspGlyTyrSerTyrGluLysGluAlaMetGluAsnThrIleSerLysLysArg 351
DB 139 TCAGATGGCTACTCTTCAAGAGAGAAAGATCAAGCTGATCCACAAAGAAAGCGT 80
QY 352 ThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsnArgThrLeu 371
DB 79 ACCAGCCCATGTGCAATTTGGCTCTTCATTCACGTGACTGACCCCAACAGACACTG 20
QY 372 LysMetAlaIleAsnArg 377
DB 19 AAGATGGCCATCAACCA 2
```

RESULT 5

BF653223/c

190 bp mRNA linear EST 25-APR-2001
276923 MARC 3BOV Bos taurus cDNA 5', mRNA sequence.

ACCESSION

BF653223

BF653223.1 GI:11918292

VERSION

EST.

Bos taurus (cow)

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 190)
Smith,T.F.L., Grose,W.M., Freking,B.A., Roberts,A.J., Stone,R.T., Casas,E., Wray,J.B., White,J., Cho,J., Fahrenkrug,S.C., Bennett,G.L., Heaton,M.P., Laegreid,W.W., Rohrer,G.A., Chitko-McKown,C.G., Perea,G., Holt,I., Karamycheva,S., Liang,F., Quackenbush,J. and Keel,J.W.

| | | | |
|---|--|---------------|-----------------------------|
| TITLE | Sequence evaluation of four pooled-tissue normalized bovine cDNA libraries and construction of a gene index for cattle | | |
| JOURNAL | Genome Res. 11 (4), 626-630 (2001) | | |
| MEDLINE | 21180013 | | |
| PUBMED | 11282978 | | |
| COMMENT | Contact: smtlch_TPL USDA, ARS, US Meat Animal Research Center PO Box 166, Clay Center, NE 68933-0166, USA Tel: 402 762 4356 Fax: 402 762 4390 Email: smtlch@mail.marc.usda.gov Single pass sequencing. Bases called and alt trimmed with phred v0.980904.e. Vector identified by cross_match with the -minscore 18 and -primers 12 options. PCR primers FORWARD: AGGAACAGCTATGACCAT BACKWARD: GTTTCACAGTCACGAC Plate: 68 row: H column: 24 Seq primer: ATTACGTGACATCATG. Location/Qualifiers 1. .190 /organism="Bos taurus" /mol_type="mRNA" /db_xref="taxon:9913" /cfeature_type="pooled" /lab_host="DH10B" /clone_id="MARCK 3BOV" /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI; library made from pooled tissue from marrow, alveolar macrophage, ovary, fetal semitendinosus muscle, and fetal longissimus muscle." | | |
| FEATURES | | | |
| SOURCE | | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Prod. NO.: | 1.68e-11 | Length: | 190 |
| Score: | 209.50 | Matches: | 43 |
| Percent Similarity: | 88.89% | Conservative: | 5 |
| Best Local Similarity: | 79.63% | Mismatches: | 4 |
| Query Match: | 10.23% | Indels: | 2 |
| DB: | 2 | Gaps: | 1 |
| US-10-077-111-13 (1-384) x BF653223 (1-190) | | | |
| Oy | 333 AepgltytYserTtYgLuLyegLuAlaMeclGluantPllEserLyLyLygkrgThr | 352 | |
| Db | 178 GAGGCTCTACTCATATGAAAGAGCAATGGAATTGATCAGCAGAAAGAACTGACC | 119 | |
| Oy | 353 SerProwctTrrAntLeuValIeuPro-SerAlaValIeuThrProAnaYgThrlLeuLy | 372 | |
| Db | 118 AGTCCCAAGACAANAATCTGTCTTCTTCAAGTGCACACTCTCCAAATAGGACTGTGA | 59 | |
| Oy | 372 sMeCa1a1eAnaYgTrPLeu--GluThrH1eGlnLyS | 384 | |
| Db | 58 AATGCCATCGATGATGGCTTAAGTGAACACACATCAAAA | 19 | |
| RESULT 6 | | | |
| LOCUS | CA587431 | 138 bp | mRNA linear EST 12-JAN-2004 |
| DEFINITION | LBR12314 cDNA from mouse aorta Mus musculus cDNA, mRNA sequence. | | |
| ACCESSION | CA587431 | | |
| VERSION | CA587431.1 GI:40792674 | | |
| KEYWORDS | EST. | | |
| SOURCE | Mus musculus (house mouse) | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| REFERENCE | 1 (bases 1 to 138) Boranz,S., Anderson,T., Thelin,A., Odeberg,J. and Lundberg,J. Vascular gene expression in atherosclerotic plaque prone regions analysed by representational difference analysis unpublished (2002) | | |
| JOURNAL | Contact: Andersson Tove Department of Biotechnology | | |
| COMMENT | | | |

TKH
Tekniskringen 34, Plan 6, 100 44 Stockholm, Sweden
Tel: +46 8 790 71 29
Fax: +46 8 245452
Email: lovebiochem.kth.se
Representations (amplified cDNA) from plaque prone regions
Seq primer: CTA TGA CCA TGA TTA CGC CAA G.
Location/Qualifiers

FEATURES
Source

1. 138
/organism="Mus musculus"
/mol_type="mRNA"
/strain="Apoe^{-/-} and LDLR^{-/-} on C57BL/6x 129 background"
/db_xref="taxon:10090"
/sex="male"
/dev_stage="8 weeks old"
/clone_1b="CDNA from mouse aorta"
/note="Organ: aorta; Site 1: DpnlI; Site 2: DpnlI, CDNA was prepared from whole aorta divided in atherosclerotic plaque prone regions (aortic arch and abdominal aorta proximal part) and less plaque prone regions (descending thoracic aorta and abdominal aorta distal part). CDNA was fragmented with DpnlI, linker ligated and amplified to generated starting material for representational difference analysis (RDA). The two cDNA pools were subjected to iterative RDA subtraction and amplification to enrich for gene fragments differentially expressed at early stages of atherosclerosis."

ORIGIN

Alignment Scores:
Pred. No.: 4.26e-08 Length: 138
Score: 174.00 Matches: 33
Percent Similarity: 91.89% Conservative: 1
Best Local Similarity: 89.19% Mismatches: 3
Query Match: 8.50% Indels: 0
DB: 6 Gaps: 0

US-10-077-111-13 (1-384) x CAG587431 (1-138)

Cy 348 LysLysLysAATgTtSerProwetThAsnLeuValLeuProSerAlaValLeuThrPro 367
|||||
Db 5 AAGAGAAAGGCTACGACCCCATGCGAAATTGGCTCTCTCCCTTCACGTAGTACGCCCA 64
|||||

Cy 368 AsnAgtThrLeuLysMetAlaIleAsnAgtTrrLeuGluThrHisGlnLys 384
|||||
Db 65 AACGAGCACTGGAAGATGCGCATCAACCGATGGCTGCGAGACGACGAGAAG 115
|||||

RESULT 7
BB323974 191 bp mRNA linear EST 11-JUL-2000
LOCUS BB323974

DEFINITION Mus musculus cDNA clone B430102G11 3', mRNA sequence.

ACCESSION BB323974

VERSION BB323974.1 GI:9032288

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murine; Mus.
1 (bases 1 to 191)

REFERENCE 1 Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T.,
Ohtsuzaki, Y., Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K.,
Shibata, K., Shibata, Y., Shigemoto, Y., Shinagawa, A., Shitaki, T.,
Sogabe, Y., Sugahara, Y., Suzuki, H., Suzuki, H., Tagawa, A.,
Takahashi, T., Tomimaga, N., Toyota, T., Tsunoda, Y., Watanabe, S., Yamamura, T., Yamane, I., Yano, R., Yasunishi, A.,
Yokota, T., Yoshida, K., Yoshiki, A., Yoshino, M., Muramatsu, M. and
Hayashizaki, Y.

TITLE
JOURNAL
COMMENT
 RIKEN Mouse ESTs (Kono, H., et al.)
 Unpublished (2000)
 Contact: Yoshihide Hayashizaki
 Laboratory for Genome Exploration Research Group, RIKEN Genomic
 Sciences Center (GSC), Yokohama Institute
 The Institute of Physical and Chemical Research (RIKEN)
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan
 Tel: 81-45-503-9222
 Fax: 81-45-503-9216
 Email: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S.,
 Saeki, N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Thermotabilization and thermocactivation of thermostable enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kitsunai, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M.,
 Okazaki, Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (<http://genome.etc.riken.go.jp/>) for
 further details.
 Location/Qualifiers
 1. .191

```

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10990"
/cldone="B430102g11"
/sex="male"
/tissue_type="adipose"
/dev_stage="4 days neonate"
/lab_host="DH10B"
/cloned_lib="Riken full-length enriched, 4 days neonate male adipose"
/notes="Site 1: Sali; Site 2: Banhi; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5]. GAGGAGAAGACATCCAAAGCGCTTTTTTTTTTTTTN 3'). cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 229.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGGAGAAGATTCGCAGGTAAATTAATAATGCCCCCCCCCCCCC 3']". cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from LambdaE.C I."
```

```

ORIGIN
Alignment Scores:
Pred. No.:      1.57e-05      length:      191
Score:          151.00      Matches:      29
Percent Similarity:      86.49%      Conservative:  2
Best Local Similarity:    78.38%      Mismatches:   5
Query Match:      7.38%      Indels:       0
DB:               2          Gaps:            0
US-10-077-111-13 (1-384) x BB323974 (1-191)

```

| | | | |
|----|-----|---|-----|
| QY | 348 | LysLysLysAaGTHSerProMetHisAsnLeuValLeuProSerAlaValLeuThrPro | 367 |
| Db | 3 | AAAGAGAAGCCTACAGACCCCATGACAAATTTGCTCTCCCTCAATTGGTACTGACCCCC | 62 |
| QY | 368 | AsnArgThrLeuLysMetAlaIleAsnArgTrrLeuGlnThrHisGlnLys | 384 |
| Db | 63 | AAAGGCACATCGAAGATGCGCTTCAACGATGGCTGGAGACCCACGAAGAAG | 113 |

[illegible]

Automated filtration-based high-throughput plasmid preparation system. Genome Res. 9 (5), 463-470 (1999)
 Carinci, P. and Hayaehizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303, 19-44 (1999)
 Please visit our web site (<http://genome.rtc.riken.go.jp>) for further details.

FEATURES

source location/Qualifiers

```
1..177
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="A330033K06"
/sex="male"
/tissue_type="spinal cord"
/dev_stage="adult"
/lab_host="DH10B"
/clone_id="RIKEN full-length enriched, adult male spinal cord"
```

/note="Site 1: SalI; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGATCCAGAGCTTTTCTTTTCTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 459.0. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGATTCGAGTTATTAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from lambda FLC I."

ORIGIN

Alignment Scores:

| | | | |
|------------------------|--------|---------------|-----|
| Pred. No.: | 0.168 | Length: | 177 |
| Score: | 111.00 | Matches: | 20 |
| Percent Similarity: | 88.46% | Conservative: | 3 |
| Best Local Similarity: | 76.92% | Mismatches: | 3 |
| Query Match: | 5.42% | Indels: | 0 |
| DB: | 2 | Gaps: | 0 |

US-10-077-111-13 (1-384) x BB186455 (1-177)

| | | | |
|----|-----|--|-----|
| QY | 358 | LeuValLeuProSerAlaValLeuThrProAsnArgThrLeuLysMetAlaIleAsnArg | 377 |
| DB | 19 | TTGCTTTCCCTTCACTACCTCCACCCCAACAGACACTGAGAGATGCCATCAACCGA | 78 |
| QY | 378 | TrpLeuGluThrHisGln | 383 |
| DB | 79 | TGCCTGGAGACCCACGAG | 96 |

Search completed: February 5, 2005, 18:04:55
 Job time : 4207 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 14:57:03 ; Search time 230 Seconds
(without alignments)
2731.871 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047
Sequence: 1 MWLIRHTLADHGDDVNCACF.....LTPNRLKAINRMLETHOK 384

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues
Total number of hits satisfying chosen parameters: 1407054

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=x1h
-Q=/cgn2_1/USFTO.spool/US10077111/runat_04022005_121141_7210/app_query.fasta_1.583
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=p2nszlm200.rn1 -MINMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bases -START=1 -END=-1 -MATRIX=blomsum62
-TRANS=human40.cdd -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200 -USER=US10077111@cgn 1.1.177@runat_04022005_121141_7210 -NCPU=6
-ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 85 | 4.2 | 191 | 4 | US-09-513-999C-21644 |
| 2 | 79.5 | 3.9 | 169 | 4 | US-09-016-434-417 |
| 3 | 66 | 3.2 | 154 | 4 | US-09-313-294A-1052 |
| 4 | 60.5 | 3.0 | 183 | 4 | US-09-902-540-8231 |
| 5 | 57 | 2.8 | 93 | 1 | US-08-190-802A-24 |
| 6 | 57 | 2.8 | 93 | 3 | US-08-477-346-24 |
| 7 | 57 | 2.8 | 93 | 3 | US-08-473-089-24 |
| 8 | 57 | 2.8 | 93 | 3 | US-08-487-072A-24 |
| 9 | 56 | 2.7 | 99 | 1 | US-08-190-802A-23 |
| 10 | 56 | 2.7 | 99 | 3 | US-08-477-346-23 |
| 11 | 56 | 2.7 | 99 | 3 | US-08-473-089-23 |
| 12 | 56 | 2.7 | 99 | 3 | US-08-487-072A-23 |

| | | | | | | | |
|---|----|------|-----|-----|---|----------------------|---------------------|
| C | 13 | 55.5 | 2.7 | 164 | 4 | US-09-513-999C-16123 | Sequence 16123, A |
| | 14 | 55.5 | 2.7 | 192 | 4 | US-09-270-767-30620 | Sequence 30620, A |
| | 15 | 55 | 2.7 | 93 | 1 | US-08-190-802A-22 | Sequence 22, Appl |
| | 16 | 55 | 2.7 | 93 | 3 | US-08-477-346-22 | Sequence 22, Appl |
| | 17 | 55 | 2.7 | 93 | 3 | US-08-473-089-22 | Sequence 22, Appl |
| | 18 | 55 | 2.7 | 93 | 3 | US-08-487-072A-22 | Sequence 22, Appl |
| | 19 | 55 | 2.7 | 151 | 4 | US-09-313-294A-1053 | Sequence 1053, Ap |
| C | 20 | 55 | 2.7 | 194 | 4 | US-09-507-765-10 | Sequence 10, Appl |
| | 21 | 54 | 2.6 | 169 | 4 | US-09-401-064-145 | Sequence 145, App |
| | 22 | 54 | 2.6 | 186 | 4 | US-09-513-999C-36540 | Sequence 36540, A |
| | 23 | 54 | 2.6 | 186 | 4 | US-09-902-540-9414 | Sequence 9414, Ap |
| | 24 | 53.5 | 2.6 | 180 | 4 | US-09-513-999C-35455 | Sequence 35455, A |
| C | 25 | 53.5 | 2.6 | 168 | 4 | US-09-902-540-4618 | Sequence 4618, Ap |
| | 26 | 53 | 2.6 | 183 | 1 | US-07-872-678A-43 | Sequence 43, Appl |
| C | 27 | 52.5 | 2.6 | 159 | 4 | US-09-270-767-3874 | Sequence 3874, Appl |
| C | 28 | 52.5 | 2.6 | 159 | 4 | US-09-270-767-19156 | Sequence 19156, A |
| C | 29 | 52.5 | 2.6 | 169 | 4 | US-09-513-999C-25155 | Sequence 25155, A |
| C | 30 | 52.5 | 2.6 | 169 | 4 | US-09-513-999C-25165 | Sequence 25165, A |
| | 31 | 52.5 | 2.6 | 195 | 4 | US-09-248-796A-8053 | Sequence 8053, Ap |
| | 32 | 52 | 2.5 | 99 | 4 | US-09-270-767-26751 | Sequence 26751, A |
| | 33 | 52 | 2.5 | 180 | 4 | US-09-513-999C-35461 | Sequence 35461, A |
| | 34 | 52 | 2.5 | 183 | 1 | US-07-872-678A-43 | Sequence 43, Appl |
| | 35 | 52 | 2.5 | 192 | 4 | US-09-248-796A-13145 | Sequence 13145, A |
| | 36 | 52 | 2.5 | 195 | 3 | US-08-717-294-2 | Sequence 2, Appl1 |
| | 37 | 52 | 2.5 | 196 | 1 | US-08-324-243-2 | Sequence 2, Appl1 |
| | 38 | 52 | 2.5 | 196 | 1 | US-08-532-390-2 | Sequence 2, Appl1 |
| | 39 | 52 | 2.5 | 196 | 5 | PCT-US95-11511-2 | Sequence 2, Appl1 |
| | 40 | 51.5 | 2.5 | 143 | 4 | US-09-513-999C-30418 | Sequence 30418, A |
| C | 41 | 51.5 | 2.5 | 176 | 4 | US-10-044-359-15 | Sequence 15, Appl |
| C | 42 | 51.5 | 2.5 | 195 | 4 | US-09-270-767-31404 | Sequence 31404, A |
| | 43 | 51 | 2.5 | 146 | 4 | US-09-468-253B-2 | Sequence 2, Appl1 |
| | 44 | 51 | 2.5 | 146 | 4 | US-09-513-999C-14749 | Sequence 14749, A |
| | 45 | 51 | 2.5 | 187 | 4 | US-09-313-294A-5562 | Sequence 5562, Ap |

ALIGNMENTS

RESULT 1
US-09-513-999C-21644

; Sequence 21644, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J. B.
; APPLICANT: Duclert, A.
; TITLE OR INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; FILE REFERENCE: 59, US2, REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26
; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 21644
; LENGTH: 191
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 136
; OTHER INFORMATION: w=a or t
US-09-513-999C-21644

Alignment Scores:

Pred. No.: 0.258
Score: 85.00
Percent Similarity: 50.00%
Best Local Similarity: 36.00%
Query Match: 4.15%
Length: 191
Matches: 18
Conservative: 7
Mismatches: 25
Indels: 0
Gaps: 0

US-10-077-111-13 (1-384) x US-09-513-999C-21644 (1-191)

[illegible]

```

Db      CTGATCAGAGCTTCTTCAGATGTACCTTAAGATTTCGACCTCTTAGAAGGAGGCTC 62
Oy      89 LeuAlaValMeGluInProSerGlySerProValArgValCysGlnPheSerProAsp 108
       :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      63 ATCTATATACCTTCAAGACATAGCGGGA---CCTGCTTTACTGTTTCATTTCAAAAGGT 119
Oy      109 SerThrCysLeuAlaSerGlyValAlaIleAspGlyThrValValLeuTrp 124
       ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db      120 GGAGAGCTATTTCATTCATCAGAGGTGCACACACACAGGCTTATTATGG 167

RESULT 3
US-09-313-294A-1052
: Sequence 1052, Application US/09313294A
: Patent No. 6476212
: GENERAL INFORMATION:
: APPLICANT: Laligudi, Raghunath V.
: APPLICANT: Ico, Laura Y.
: APPLICANT: Sherman, Bradley K.
: TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN BAR
: FILE REFERENCE: PL-0017 US
: CURRENT APPLICATION NUMBER: US/09/313,294A
: PRIOR FILING DATE: 1999-05-14
: NUMBER OF SEQ ID NOS: 7600
: SOFTWARE: PERL Program
: SEQ ID NO 1052
: LENGTH: 154
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No. 6476212 700550309H1
US-09-313-294A-1052

Alignment Scores:
Pred. No.:
Score: 36.7 Length: 154
Percent Similarity: 66.00 Matches: 13
Best Local Similarity: 43.33% Conservative: 5
Query Match: 3.22% Mismatches: 12
DB: 4 Indels: 0
Gaps: 0

US-10-077-111-13 (1-384) x US-09-313-294A-1052 (1-154)

Oy      99 ProValArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAsp 118
       :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db      38 CCCTGCTCTGTGTTGATTTCAGTGGAGATCTCGAAATGCTGGCATCTGATCAGAGAT 97
Oy      119 GlyThrValValLeuTrpPheAlaIleSer 128
       ||| :::: ||||| ::::
Db      98 GGAAAGATTAGGTTGGCGCTATACGAAC 127

RESULT 4
US-09-902-540-8231/c
: Sequence 8231, Application US/09902540
: Patent No. 6833447
: GENERAL INFORMATION:
: APPLICANT: Goldman, Barry S.
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Slater, Steven C.
: APPLICANT: Wiegand, Roger C.
: TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
: FILE REFERENCE: 38-10(15849)B
: CURRENT APPLICATION NUMBER: US/09/902,540
: PRIOR FILING DATE: 2001-07-10
: PRIOR APPLICATION NUMBER: 60/217,863
: PRIOR FILING DATE: 2000-07-10
: NUMBER OF SEQ ID NOS: 16825
: SEQ ID NO 8231
: LENGTH: 163
: TYPE: DNA
: ORGANISM: Myxococcus xanthus
US-09-902-540-8231

```


Alignment Scores:

Score: 232
Pred. No.: 173
Percent Similarity: 53.66%
Best Local Similarity: 41.46%
Query Match: 2.96%
DB: 4
Length: 183
Matches: 17
Conservative: 5
Mismatch: 10
Indels: 9
Gaps: 2

US-10-077-111-13 (1-384) x US-09-902-540-8231 (1-183)

QY 345 TYPILSErlySylsYsArGThrSerPROMeThrAsnLeuValLeuProSerAlaVal 364

Db 173 TGGATGTCGTGATGAAAGAACGACG-----CGAGTCGTCTTCATCGCGCGGTC 123

QY 365 -----LeuThrProAsnArGThrLeuLysMetAlaIleAsnArGTrp 378

Db 122 CATCCGAGCCACCACCACTGTCCATACCGAGCGCGGTGATGCTGACTGCG 63

QY 379 Leu 379

Db 62 CTC 60

RESULT 5

US-08-190-802A-24

Sequence 24, Application US/08190802A

Patent No. 5519003

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theoreof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Dehlinger & Associates

STREET: P.O. Box 60850

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94306-0850

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/190.802A

FILING DATE: 01-FEB-1994

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: Fabian, Gary R.

REGISTRATION NUMBER: 33,875

REFERENCE/DOCKET NUMBER: 8600-0139

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 324-0880

TELEFAX: (415) 324-0960

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 93 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: RACK1 rv DNA Sequence, Fig. 1A

US-08-190-802A-24

Alignment Scores:

Pred. No.: 191

Score: 57.00

Percent Similarity: 65.00%

Best Local Similarity: 50.00%

Length: 93

Matches: 10

Conservative: 3

Mismatch: 7

Best Local Similarity: 50.00%

Query Match:

DB: 1
Indels: 0
Gaps: 0

US-10-077-111-13 (1-384) x US-08-190-802A-24 (1-93)

QY 106 SerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTPAsn 125

Db 34 TCTCCAGATGATCCCTCTGTGCTTCTGAGCGAGAGATGCGCAGGCTATGCTGTGGAT 93

RESULT 6

US-08-477-346-24

Sequence 24, Application US/08477346

Patent No. 6262023

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theoreof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESS: Morrison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/477,346

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/487,072

FILING DATE: 07-JUN-1995

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 24:

SEQUENCE CHARACTERISTICS:

LENGTH: 93 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: RACK1 rv DNA Sequence, Fig. 1A

US-08-477-346-24

Alignment Scores:

Pred. No.: 191

Score: 57.00

Percent Similarity: 65.00%

Best Local Similarity: 50.00%

Length: 93

Matches: 10

Conservative: 3

Mismatch: 7

Query Match: 2.78%

Indels: 0

Gaps: 0

DB: 3

US-10-077-111-13 (1-384) x US-08-477-346-24 (1-93)

QY 106 SerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTPAsn 125

Db 34 TCTCCAGATGATCCCTCTGTGCTTCTGAGCGAGAGATGCGCAGGCTATGCTGTGGAT 93

RESULT 7

```

US-08-473-089-24
; Sequence 24, Application US/08473089
; Patent No. 63423568
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,089
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ. ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEetical: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 rV DNA sequence, Fig. 1A
US-08-473-089-24

Alignment Scores:
Pred. No.: 191 Length: 93
Score: 57.00 Matches: 10
Percent Similarity: 65.00% Conservative: 3
Best Local Similarity: 50.00% Mismatches: 7
Query Match: 2.78% Indels: 0
DB: 3 Gaps: 0

US-10-077-111-13 (1-364) x US-08-473-089-24 (1-93)

QY 106 SerProAspSerThrcysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTTPAsn 1255
Db 34 TCTCCAGATGATCCCTCTGTGCTCTTGAGGCGCAAGATGCGCAGGCTATGCTGTGGAT 93
:::|||||:::|||||:::|||||:::

```

```

1 STATE: DC
2 COUNTRY: USA
3 ZIP: 20006-1812
4
5 COMPUTER READABLE FORM:
6
7 MEDIUM TYPE: Floppy disk
8 COMPUTER: IBM PC compatible
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10 SOFTWARE: Patentin Release #1.0, Version #1.25
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/08/487,072A
13 FILING DATE: 07-JUN-1995
14 CLASSIFICATION: 514
15
16 ATTORNEY/AGENT INFORMATION:
17 NAME: MURASHIGE, KATE H.
18 REGISTRATION NUMBER: 29,959
19 REFERENCE/DOCKET NUMBER: 2550-0025.20
20 TELECOMMUNICATION INFORMATION:
21 TELEPHONE: (202) 887-1500
22 TELEFAX: (202) 887-0763
23 INFORMATION FOR SEQ ID NO: 24:
24 SEQUENCE CHARACTERISTICS:
25 LENGTH: 93 base pairs
26 TYPE: nucleic acid
27 STRANDEDNESS: double
28 TOPOLOGY: linear
29 MOLECULE TYPE: DNA (genomic)
30
31 HYPOTHEetical: NO
32 ANTI-SENSE: NO
33 ORIGINAL SOURCE:
34 INDIVIDUAL ISOLATE: RACK1 rV DNA sequence, Fig. 1A
35
36 US-08-487-072A-24
37
38 Alignment Scores:
39 Pred. No.: 191 Length: 93
40 Score: 57.00 Matches: 10
41 Percent Similarity: 65.00% Conservative: 3
42 Best local Similarity: 50.00% Mismatches: 7
43 Query Match: 2.78% Indels: 0
44 DB: 3 Gaps: 0
45
46 US-10-077-111-13 (1-384) x US-08-487-072A-24 (1-93)
47
48 Oy 106 SerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValValLeuTrpAsn 125
49 Db 34 TCCTCGAGATGATCCTCTGTGCTCTTGTGAGGCAAGATGGCCAGGCTATGCTGTGGAT 93
50
51 RESULT 9
52 US-08-190-802A-23
53 Sequence 23, Application US/08190802A
54 Patent No. 5519003
55
56 GENERAL INFORMATION:
57 APPLICANT: Mochly-Rosen, Daria
58 APPLICANT: Ron, Dorit
59 TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
60 TITLE OF INVENTION: Thereof
61 NUMBER OF SEQUENCES: 265
62 CORRESPONDENCE ADDRESS:
63 ADDRESSEE: Dehlinger & Associates
64 STREET: P.O. Box 60850
65 CITY: Palo Alto
66 STATE: CA
67 COUNTRY: USA
68 ZIP: 94306-0850
69
70 COMPUTER READABLE FORM:
71 MEDIUM TYPE: Floppy disk
72 COMPUTER: IBM PC compatible
73 OPERATING SYSTEM: PC-DOS/MS-DOS
74 SOFTWARE: Patentin Release #1.0, Version #1.25
75 CURRENT APPLICATION DATA:
76 APPLICATION NUMBER: US/08/190,802A
77 FILING DATE: 01-FEB-1994
78 CLASSIFICATION: 530
79 ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEITICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 r1v DNA Sequence, Fig. 1A
; US-08-190-802A-23

Alignment Scores:
Pred. No.: 282 Length: 99
Score: 56.00 Matches: 13
Percent Similarity: 50.00% Conservative: 3
Best Local Similarity: 40.62% Mismatches: 14
Query Match: 2.74% Indels: 2
DB: 1 Gaps: 1

US-10-077-111-13 (1-384) x US-08-190-802A-23 (1-99)

QY 53 HisThrTyraIaValHisCysCysePhseSerPro-----SergIyHisIleLeuAla 70
Db 4 CATTCAGATGGGTGCTGTGTCGCGCTTCCCCGAACAGACAGAACCCATCATGTC 63

QY 71 SerCysSerThraPrgIyThrThrValLeuTrpAsn 82
Db 64 TCCTGGAGATGGGACCAAGCTGTCAAGGTGTGGAT 99

RESULT 10
US-08-477-346-23
; Sequence 23, Application US/08477346
; Patent No. 6262023
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/477,346
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/487,072
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
```

```

; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEITICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 r1v DNA Sequence, Fig. 1A
; US-08-477-346-23

Alignment Scores:
Pred. No.: 282 Length: 99
Score: 56.00 Matches: 13
Percent Similarity: 50.00% Conservative: 3
Best Local Similarity: 40.62% Mismatches: 14
Query Match: 2.74% Indels: 2
DB: 3 Gaps: 1

US-10-077-111-13 (1-384) x US-08-477-346-23 (1-99)

QY 53 HisThrTyraIaValHisCysCysePhseSerPro-----SergIyHisIleLeuAla 70
Db 4 CATTCAGATGGGTGCTGTGTCGCGCTTCCCCGAACAGACAGAACCCATCATGTC 63

QY 71 SerCysSerThraPrgIyThrThrValLeuTrpAsn 82
Db 64 TCCTGGAGATGGGACCAAGCTGTCAAGGTGTGGAT 99

RESULT 11
US-08-473-089-23
; Sequence 23, Application US/08473089
; Patent No. 6342368
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morrison & Foerster
; STREET: 2000 Pennsylvania Avenue, NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20006-1812
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/473,089
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MURASHIGE, KATE H.
; REGISTRATION NUMBER: 29,959
; REFERENCE/DOCKET NUMBER: 2550-0025.22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 887-1500
; TELEFAX: (202) 887-0763
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 99 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEITICAL: NO
; ANTI-SENSE: NO
```

ORIGINAL SOURCE: RACK1 r1v DNA Sequence, Fig. 1A
INDIVIDUAL ISOLATE: US-08-473-089-23

Alignment Scores:

| Pred. No.: | 282 | Length: | 99 |
|------------------------|--------|---------------|----|
| Score: | 56.00 | Matches: | 13 |
| Percent Similarity: | 50.00% | Conservative: | 3 |
| Best Local Similarity: | 40.62% | Mismatches: | 14 |
| Query Match: | 2.74% | Indels: | 2 |
| DB: | 3 | Gaps: | 1 |

US-10-077-111-13 (1-384) x US-08-473-089-23 (1-99)

RESULT 12

US-08-487-072A-23

Sequence 23, Application US/08487072A

Patent No. 6423684

GENERAL INFORMATION:

APPLICANT: Mochly-Rosen, Daria

APPLICANT: Ron, Dorit

TITLE OF INVENTION: WD-40 - Derived Peptides and Uses

TITLE OF INVENTION: Theof

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Morison & Foerster

STREET: 2000 Pennsylvania Avenue, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1812

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent in Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,072A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: MURASHIGE, KATE H.

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 2550-0025.20

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 887-1500

TELEFAX: (202) 887-0763

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 99 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

HYPOTHETICAL: NO

ANTI-SENSE: NO

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: RACK1 r1v DNA Sequence, Fig. 1A

US-08-487-072A-23

Alignment Scores:

| Pred. No.: | 282 | Length: | 99 |
|------------------------|--------|---------------|----|
| Score: | 56.00 | Matches: | 13 |
| Percent Similarity: | 50.00% | Conservative: | 3 |
| Best Local Similarity: | 40.62% | Mismatches: | 14 |

Query Match: 2.74% Indels: 2
DB: 3 Gaps: 1

US-10-077-111-13 (1-384) x US-08-487-072A-23 (1-99)

Query 53 HistHrTyrAlaValHisCysCysPheSerPro-----SerGlyHisIleLeuAla 70
|||||
4 CATTCGAATGGGTGCTTGTGTCGCGTTCTCCCGAACAAGCAACCTATCATGTC 63

Query 71 SerCysSerThrAaspGlyThrValLeuTPAen 82
|||||
64 TCCTGGGATGGAGCAAGCTGCTCAAGTGTGGAT 99

RESULT 13

US-09-513-999C-16123/C

Sequence 16123, Application US/09513999C

Patent No. 6783961

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Duclert, A.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

Patent No. 6783961

FILE REFERENCE: 59.US2.REG

CURRENT APPLICATION NUMBER: US/09/513,999C

CURRENT FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent.pm

SEQ ID NO 16123

LENGTH: 164

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: msc_feature

LOCATION: 12

OTHER INFORMATION: y=c or t

US-09-513-999C-16123

Alignment Scores:

| Pred. No.: | 779 | Length: | 164 |
|------------------------|--------|---------------|-----|
| Score: | 55.50 | Matches: | 16 |
| Percent Similarity: | 48.15% | Conservative: | 10 |
| Best Local Similarity: | 29.63% | Mismatches: | 19 |
| Query Match: | 2.71% | Indels: | 9 |
| DB: | 4 | Gaps: | 2 |

US-10-077-111-13 (1-384) x US-09-513-999C-16123 (1-164)

Query 12 GLYAspAspValAsnCysCys-----AlaPheSerPheSerLeuLeuAlaThrCysSer 29
|||||
154 GGAGACCAAGTCTCGCTGTCTATCAGCTGGAGTGCAGCAATCATCTACTGACG 95

Query 30 LeuAspLysThrIleArgLeuTyrSerLeuAArgAspPheThrGluLeuProHisSerPro 49
|||||
94 CTGGAA-----CTCCAAAGCCCAAGGATGCCCACTCAGCT 56

Query 50 LeuLysPheHisThrTyrAlaValHisCysCysPheSer 63
|||||
55 TCCAAAGTGTAACTATAGTGTGAGCTACCATGCTGGCC 14

RESULT 14

US-09-270-767-30620

Sequence 30620, Application US/09270767

Patent No. 6703491

GENERAL INFORMATION:

APPLICANT: Homburger et al.

TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

FILE REFERENCE: File Reference: 7326-094

CURRENT APPLICATION NUMBER: US/09/270,767

CURRENT FILING DATE: 1999-03-17

NUMBER OF SEQ ID NOS: 62517

```
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30620
; LENGTH: 192
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-30620

Alignment Scores:
Pred. No.: 1.02e+03 Length: 192
Score: 55.50 Matches: 14
Percent Similarity: 47.06% Conservative: 10
Best Local Similarity: 27.45% Mismatches: 26
Query Match: 2.71% Indels: 1
DB: 4 Gaps: 1

US-10-077-111-13 (1-384) x US-09-270-767-30620 (1-192)

QY 230 GluH1eGlnLeuYsgInPheThrGluAspTrpSerGluGluValSerThrTrpLeu 249
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 19 GAGCCCTGGCTCAGATCGAGAGTGCATTAATTGACATCGACGACACCGATTGCTG 78

QY 250 Cye---AlaGlnAspLeuYAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
   : : : : : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 GCTCAGTCCGCTTCAGCTGCCGCAATACGTTGATCTGTCAATTACACAAAGTTACTGCGC 138

QY 269 LysGluLeuLeuAsnLeuThrIlyGluSerIleu 279
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 GCTGCCCTTGGCCAGATTGGCTGTGATTAATCTT 171

RESULT 15
US-08-190-802A-22
; Sequence 22, Application US/08190802A
; Patent No 5519003
; GENERAL INFORMATION:
; APPLICANT: Mochly-Rosen, Daria
; APPLICANT: Ron, Dorit
; TITLE OF INVENTION: WD-40 - Derived Peptides and Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dehlinger & Associates
; STREET: P.O. Box 60850
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306-0850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/190,802A
; FILING DATE: 01-FEB-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Fabian, Gary R.
; REGISTRATION NUMBER: 33,875
; REFERENCE/DOCKET NUMBER: 8600-0139
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 324-0880
; TELEFAX: (415) 324-0960
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 93 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: RACK1 r111 DNA Sequence, Fig. 1A
```

```
US-08-190-802A-22

Alignment Scores:
Pred. No.: 335 Length: 93
Score: 55.00 Matches: 10
Percent Similarity: 66.67% Conservative: 4
Best Local Similarity: 47.62% Mismatches: 7
Query Match: 2.69% Indels: 0
DB: 1 Gaps: 0

US-10-077-111-13 (1-384) x US-08-190-802A-22 (1-93)

QY 105 PheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValLeuTrp 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 31 TTCTCCTGTGACACCGGACGAGATTGTCTGTGGTCCCGAGACAAACCATTAATTATGG 90

QY 125 Asn 125
   |||
Db 91 AAT 93

Search completed: February 5, 2005, 18:08:50
Job time : 232 secs
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 04:18:44 ; Search time 763 Seconds
(without alignments)
9868.822 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gaattcggttcactctgcg.....acaccaaagtaagaatcc 1272

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_16Dec04:*
1: geneseqn1980s:*
2: geneseqn1980s:*
3: geneseqn2000s:*
4: geneseqn2001s:*
5: geneseqn2001s:*
6: geneseqn2002s:*
7: geneseqn2002s:*
8: geneseqn2003s:*
9: geneseqn2003s:*
10: geneseqn2003s:*
11: geneseqn2003s:*
12: geneseqn2004s:*
13: geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|---------------------|
| 1 | 1272 | 100.0 | 1272 | 6 | AAD45075 Human RET |
| 2 | 977.2 | 76.8 | 1553 | 6 | AAD45071 Human RET |
| 3 | 965.8 | 75.9 | 1818 | 6 | AAD45070 Human RET |
| 4 | 964.2 | 75.8 | 1817 | 10 | ADB47507 Human CDN |
| 5 | 964.2 | 75.8 | 1844 | 4 | AA158876 Human pol |
| 6 | 964.2 | 75.8 | 1844 | 5 | ADQ99098 DNA encod |
| 7 | 964.2 | 75.8 | 1844 | 9 | ADB48858 Novel hum |
| 8 | 894 | 70.3 | 1773 | 8 | ABZ24707 Human cel |
| 9 | 890.2 | 70.0 | 1811 | 10 | ADB63535 Human CDN |
| 10 | 878.8 | 69.1 | 1996 | 5 | ABV23028 Human pro |
| 11 | 878.8 | 69.1 | 1996 | 5 | ABV23190 Human pro |
| 12 | 877.8 | 69.0 | 1908 | 6 | AAD45076 Human RET |
| 13 | 867.8 | 68.2 | 1826 | 4 | AA160662 Human pol |
| 14 | 791 | 62.2 | 1623 | 5 | AA164914 Beta-rian |
| 15 | 782 | 61.5 | 1686 | 4 | ADM19385 Novel hum |
| 16 | 654 | 51.4 | 838 | 5 | ADM19628 Novel hum |
| 17 | 478.6 | 37.6 | 1291 | 4 | AAFS58364 Human GTP |
| 18 | 464 | 36.5 | 1901 | 6 | AAD45073 Mouse RET |
| 19 | 390 | 30.7 | 668 | 13 | ADQ54461 Novel can |
| 20 | 377.2 | 29.7 | 630 | 6 | AAD45072 Human RET |

| | | | | | | |
|---|----|-------|------|-----|---|--------------------|
| C | 21 | 335 | 26.3 | 446 | 5 | ABV05038 Human pro |
| C | 22 | 331.6 | 26.1 | 366 | 4 | AA136816 Probe #55 |
| C | 23 | 331.6 | 26.1 | 366 | 4 | AAK30903 Human bon |
| C | 24 | 331.6 | 26.1 | 366 | 4 | AAK05311 Human bra |
| C | 25 | 331.6 | 26.1 | 366 | 4 | ABS30581 Human liv |
| C | 26 | 331.6 | 26.1 | 366 | 6 | ABS05651 Human gen |
| C | 27 | 329.8 | 25.9 | 409 | 5 | ABV14207 Human pro |
| C | 28 | 328.8 | 25.8 | 441 | 5 | ABV35301 Human pro |
| C | 29 | 328.8 | 25.8 | 441 | 5 | ABV44135 Human pro |
| C | 30 | 297 | 23.3 | 297 | 4 | AA150015 Probe #18 |
| C | 31 | 297 | 23.3 | 297 | 4 | AAK44005 Human bon |
| C | 32 | 297 | 23.3 | 297 | 4 | AAK18116 Human bra |
| C | 33 | 297 | 23.3 | 297 | 4 | ABS43659 Human liv |
| C | 34 | 297 | 23.3 | 297 | 6 | ABS18238 Human gen |
| C | 35 | 291.6 | 22.9 | 419 | 3 | AAH30188 Human col |
| C | 36 | 179.6 | 14.1 | 466 | 4 | ABA58107 Human foe |
| C | 37 | 179.6 | 14.1 | 466 | 4 | AA137710 Probe #63 |
| C | 38 | 179.6 | 14.1 | 466 | 4 | AAK31840 Human bon |
| C | 39 | 179.6 | 14.1 | 466 | 4 | AAK06182 Human bra |
| C | 40 | 179.6 | 14.1 | 466 | 4 | ABS31528 Human liv |
| C | 41 | 179.6 | 14.1 | 466 | 6 | ABS06599 Human gen |
| C | 42 | 176.8 | 13.9 | 180 | 4 | ABA70710 Human foe |
| C | 43 | 176.8 | 13.9 | 180 | 4 | AA150879 Probe #19 |
| C | 44 | 176.8 | 13.9 | 180 | 4 | AAK4908 Human bon |
| C | 45 | 176.8 | 13.9 | 180 | 4 | AAK18959 Human bra |

ALIGNMENTS

| | | |
|----------|---|--|
| RESULT 1 | AAAD45075 | standard; DNA; 1272 BP. |
| ID | AAAD45075 | |
| AC | AAAD45075; | |
| DT | 27-DEC-2002 | (first entry) |
| DE | Human RET16.2 splice variant DNA. | |
| XX | Human; RET16; intracellular signal; inflammation-related disease; asthma; | |
| KW | rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective; | |
| KW | transplant rejection; chronic obstructive pulmonary disease; TNF-alpha; | |
| KW | inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer; | |
| KW | inacute respiratory distress syndrome; candida; ulcerative colitis; | |
| KW | autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm; | |
| KW | cellular migration disorder; cell proliferation disorder; calcification; | |
| KW | hyperinsulinaemia; diabetes type 2; systemic lupus erythematosus; tumour; | |
| KW | cardiovascular disease; Wegener's granulomatosis; atherosclerosis; | |
| KW | thalassaemia; vasotropic; gene; ds. | |
| XX | Homo sapiens. | |
| OS | | |
| XX | | |
| FX | Key | Location/Qualifiers |
| FT | CDS | 111..1265 |
| FT | | /*tag= a |
| FT | | /product= "Human RET16.2 splice variant protein" |
| XX | WO200266494-A2. | |
| PD | 29-AUG-2002. | |
| XX | | |
| PF | 15-FEB-2002; 2002WO-US005162. | |
| XX | | |
| PR | 16-FEB-2001; 2001US-0269366P. | |
| PR | 29-MAY-2001; 2001US-0294181P. | |
| XX | | |
| PA | (BRIM) BRISTOL-MYERS SQUIBB CO. | |
| XX | | |
| PI | Toddernud CG, Finger JN, Rillema J; | |
| XX | | |
| DR | WPI; 2002-682760/73. | |
| DR | P-PSDB; AAE28167. | |

XX New human, mouse or rat RET16 genes and proteins, involved in
PT intracellular signaling cascade, useful for in gene therapy, particularly
PT for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
PT tumors or neoplasms.

PS Claim 1, Page 166; 175pp; English.

CC The invention relates to human mouse or rat Rtl16 genes and proteins,
CC involved in intracellular signaling cascade. The Rtl16 protein or
CC polynucleotide is useful for treating an inflammation-related disease or
CC disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis,
CC asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or
CC tissue transplants, chronic obstructive pulmonary disease, inflammatory
CC bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory
CC distress syndrome, systemic lupus erythematosus, autoimmune disease,
CC cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
CC disease or disorder also includes disorders associated with aberrant
CC activation of the TNF- α pathway, disorders associated with aberrant
CC cellular migration, proliferation, metastasis, juvenile idiopathic
CC arthritis, haematogenous metastases of tumour cells, hypernucleaemia,
CC diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer,
CC tumour progression, Megener's granulomatosis, stem cell transplantation
CC complications, ischaemia-reperfusion injury, thalasassaemia, acute lung
CC injury, graft rejection, ischaemic heart, coronary artery calcification
CC or allergic inflammation. Rtl16 DNA is used in gene therapy. The present
CC sequence is human Rtl16.2 splice variant DNA

Sequence 1272 BP; 349 A; 269 C; 290 G; 364 T; 0 U; 0 Other;

| Query Match | Score | DB | Length |
|-------------|-------|----|--------|
| 100.0% | 1272 | 6 | 1272 |

Matches 1272; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

| | | | | |
|----|-----|--|-------------------------------|----|
| QY | 1 | GAATTGGAGCTTTTCACTCGGCGGACAGTGA | CCCCGACCCGCGGTGGGACCTTTGAAGCG | 60 |
| Db | 1 | GAATTGGAGCTTTTCACTCGGCGGACAGTGA | CCCCGACCCGCGGTGGGACCTTTGAAGCG | 60 |
| QY | 61 | GATCCCGGCGCGCGCGCTCTCGCAGGCTGTTTTCTTCAATAAAGAAATGCTGAAC | 120 | |
| Db | 61 | GATCCCGGCGCGCGCGCTCTCGCAGGCTGTTTTCTTCAATAAAGAAATGCTGAAC | 120 | |
| QY | 121 | TGATTCACACATTAGCTGATTCATGATGACGATGTCAACTGCTGCGCTTCTCCTTTTCCC | 180 | |
| Db | 121 | TGATTCACACATTAGCTGATTCATGATGACGATGTCAACTGCTGCGCTTCTCCTTTTCCC | 180 | |
| QY | 181 | TCTTGGCTACTTGGCTCTCTTGGACAAAACAATTGGCTGTACTGTGTTACGTGACTTTACTG | 240 | |
| Db | 181 | TCTTGGCTACTTGGCTCTCTTGGACAAAACAATTGGCTGTACTGTGTTACGTGACTTTACTG | 240 | |
| QY | 241 | AACTGCGACATTTCTCCATTGAAGTTTCATPCTATGCTGTGCCACTGCTGCTGTTCTCCC | 300 | |
| Db | 241 | AACTGCGACATTTCTCCATTGAAGTTTCATPCTATGCTGTGCCACTGCTGCTGTTCTCCC | 300 | |
| QY | 301 | CTTCAGGACATATTTTGGCATCGTGTTCACAAGATGGTACCACTGTCTTATGGAATACTG | 360 | |
| Db | 301 | CTTCAGGACATATTTTGGCATCGTGTTCACAAGATGGTACCACTGTCTTATGGAATACTG | 360 | |
| QY | 361 | AAAAATGACAGATGCTGCGCAGTATGGAACAGCTTACTGTGGCAGCCCTGTGAGGCTTTGCC | 420 | |
| Db | 361 | AAAAATGACAGATGCTGCGCAGTATGGAACAGCTTACTGTGGCAGCCCTGTGAGGCTTTGCC | 420 | |
| QY | 421 | AGTTTTCCCCAGCTCAAGTGTTTGGCATCAGGGGAGCTGATGGAACCTGTGGTTTTGT | 480 | |
| Db | 421 | AGTTTTCCCCAGCTCAAGTGTTTGGCATCAGGGGAGCTGATGGAACCTGTGGTTTTGT | 480 | |
| QY | 481 | GGAAATGACAGTCAATCAAAATTATATAGATGTGTATGTTAAAGATGCTCCTTGGCGG | 540 | |
| Db | 481 | GGAAATGACAGTCAATCAAAATTATATAGATGTGTATGTTAAAGATGCTCCTTGGCGG | 540 | |
| QY | 541 | CATGTGCAATTTTCTCTTAATGGAAGCTCTTTTGTCACTGGCTCTCATGTGTGATTTAA | 600 | |
| Db | 541 | CATGTGCAATTTTCTCTTAATGGAAGCTCTTTTGTCACTGGCTCTCATGTGTGATTTAA | 600 | |

| | | | |
|----------|---|--|------|
| Qy | 601 | CAGGTGGGATGATTAATAATGAGGTGTCTGCATAGTGAATAAAGCATGATCTTGGAAATTA | 660 |
| Db | 601 | CAGGTGGGATGATTAATAATGAGGTGTCTGCATAGTGAATAAAGCATGATCTTGGAAATTA | 660 |
| Qy | 661 | CTGTGTGGATTTTTCCTTCACAGCGAGTTCTGATGTGAGAAACAAGTCTTCAGTTTTC | 720 |
| Db | 661 | CTGTGTGGATTTTTCCTTCACAGCGAGTTCTGATGTGAGAAACAAGTCTTCAGTTTTC | 720 |
| Qy | 721 | GACTGGCATCATGTGTTCAGGATTTGCCAAGTCGAAATTTGGATTGTTCTTTTACCATTA | 780 |
| Db | 721 | GACTGGCATCATGTGTTCAGGATTTGCCAAGTCGAAATTTGGATTGTTCTTTTACCATTA | 780 |
| Qy | 781 | TCTTAAGCAAGGCGACAGAACATAGCTGAGAAACAATTAACGAATTTGGTCAGAGAGG | 840 |
| Db | 781 | TCTTAAGCAAGGCGACAGAACATAGCTGAGAAACAATTAACGAATTTGGTCAGAGAGG | 840 |
| Qy | 841 | TGCTCTCAACATGGCTTTGTGTCACAAAGATTTTAAAGATCTTGTTGATATTTTCACAATGA | 900 |
| Db | 841 | TGCTCTCAACATGGCTTTGTGTCACAAAGATTTTAAAGATCTTGTTGATATTTTCACAATGA | 900 |
| Qy | 901 | ATACATTTGATGAGAAAGAACTGTTGAATCTTACAAAAGAAAGCTGTGGCTGATATTTGA | 960 |
| Db | 901 | ATACATTTGATGAGAAAGAACTGTTGAATCTTACAAAAGAAAGCTGTGGCTGATATTTGA | 960 |
| Qy | 961 | AAATTGAATCTCTAGGACTGGCTGATGAAGTGCAGAGAAATTTGAAGAGCTCAGACCA | 1020 |
| Db | 961 | AAATTGAATCTCTAGGACTGGCTGATGAAGTGCAGAGAAATTTGAAGAGCTCAGACCA | 1020 |
| Qy | 1021 | AGGTTAAATCCCTTCTTCACAGAAATTCCTGTGAATTAATGTCATTAATTAAGAGAAC | 1080 |
| Db | 1021 | AGGTTAAATCCCTTCTTCACAGAAATTCCTGTGAATTAATGTCATTAATTAAGAGAAC | 1080 |
| Qy | 1081 | TTATGAAAGATCCGGTCACTGCCATCAGATGGCTATTATATGAAAAGAAAGCATGAGAA | 1140 |
| Db | 1081 | TTATGAAAGATCCGGTCACTGCCATCAGATGGCTATTATATGAAAAGAAAGCATGAGAA | 1140 |
| Qy | 1141 | ATTGATCAGAGAAAGAAAGGTAACAAGTCCCATGACAAATCTTGTTCTTCCACACGG | 1200 |
| Db | 1141 | ATTGATCAGAGAAAGAAAGGTAACAAGTCCCATGACAAATCTTGTTCTTCCACACGG | 1200 |
| Qy | 1201 | TACTTACACCAATAGAGCTGTGAAATATGSCCATCAATAGATGGCTGAGACACACCAA | 1260 |
| Db | 1201 | TACTTACACCAATAGAGCTGTGAAATATGSCCATCAATAGATGGCTGAGACACACCAA | 1260 |
| Qy | 1261 | AGTAAGAATTC 1272 | |
| Db | 1261 | AGTAAGAATTC 1272 | |
| RESULT 2 | | | |
| AAD45071 | | | |
| ID | AAD45071 standard; DNA, 1553 BP. | | |
| AC | AAD45071; | | |
| XX | | | |
| XX | 27-DEC-2002 (first entry) | | |
| DE | Human RET16 open reading frame (ORF) DNA. | | |
| XX | | | |
| XX | Human; RET16; intracellular signal; inflammation-related disease; asthma; | | |
| KW | rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective; | | |
| KW | transplant rejection; chronic obstructive pulmonary disease; TNF-alpha; | | |
| KW | inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer; | | |
| KW | inacute respiratory distress syndrome; cardiac; ulcerative colitis; | | |
| KW | autoimmune disease; cystic fibrosis; gene therapy; cytotoxic; neoplasm; | | |
| KW | cellular migration disorder; cell proliferation disorder; calcification; | | |
| KW | hyperinsulinaemia; diabetes type 2; systemic lupus erythematosus; tumour; | | |
| KW | cardiovascular disease; Wegener's granulomatosis; atherosclerosis; | | |
| KW | thalassaemia; vasotrophic; open reading frame; ORF; gene; de. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |

| | | | |
|----|---|--|--|
| FH | Key | | Location/Qualifiers |
| FT | CDS | 3..1541 | /tag=a /product= "Human RETL6 protein" /note= "CDS does not include start codon" |
| FT | | | /partial |
| XX | | | |
| PX | | | |
| PA | (BRIM) BRISTOL-MYERS SQUIBB CO. | | |
| XX | | | |
| XI | Toderdud CG, Finger JN, Rillema U; | | |
| DV | WPI; 2002-682760/73. | | |
| DR | P-PSTDB; AAE28164. | | |
| PT | New human, mouse or rat RETL6 genes and proteins, involved in intracellular signaling cascade, useful for in gene therapy, particularly for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers, tumors or neoplasms. | | |
| PS | Claim 1; Page 150; 175pp; English. | | |
| CC | The invention relates to human, mouse or rat RETL6 genes and proteins, | | |
| CC | involved in intracellular signaling cascade. The RETL6 protein or polynucleotide is useful for creating an inflammation-related disease or disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis, ashma, ischemia-reperfusion, multiple sclerosis, rejection of organ or tissue transplants, chronic obstructive pulmonary disease, inflammatory bowel disease, Crohn's disease, ulcerative colitis, inactive respiratory distress syndrome, systemic lupus erythematosus, autoimmune disease, cyclic fibroses, cancers, tumours or neoplasms. This inflammation-related disease or disorder also includes disorders associated with aberrant activation of the TNF-alpha pathway, disorders associated with aberrant cellular migration, proliferation, metastasis, juvenile idiopathic arthritis, haematogenous metastases of tumour cells, hyperinulinaemia, diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer, tumour progression, Wegener's granulomatosis, stem cell transplantation complications, ischaemic-reperfusion injury, thalassemia, acute lung injury, graft rejection, ischaemic heart, coronary artery calcification or allergic inflammation. RETL6 DNA is used in gene therapy. The present sequence is human RETL6 open reading frame (ORF) DNA | | |
| SQ | Sequence 1553 BP; 428 A; 327 C; 348 G; 450 T; 0 U; 0 Other; | | |
| | Query Match | 76.8%; Score 977.2; DB 6; Length 1553; | |
| | Best Local Similarity | 81.9%; Pred. No. 4,9e-288; | |
| | Matches 1265; Conservative | 0; Mismatches 3; Indels 276; Gaps 1 | |
| OY | 1 GAATTCGGCTTTCACTCGCGGAGCATGTGAACGCCGCCTTGGGACTTTGAAGCG | 60 | |
| Db | 1 GAATTGGCGCTTTCACTCGCGGAGCATGTGAACGCCGCCTTGGGACTTTGAAGCG | 60 | |
| OY | 61 GATCCCGGCGCCCCCGCTCTGTGAAGCGTGTTTTCTTAAATAAAGAACAATGTGAAC | 120 | |
| Db | 61 GATCCCGGCGCCCCCGCTCTGTGAAGCGTGTTTTCTTAAATAAAGAACAATGTGAAC | 120 | |
| OY | 121 TGATTCACAANTTAAGTAGATGAGTGAAGAGTGAACGTGCTGCTTCCCTTTCCC | 180 | |
| Db | 121 TGATTCACAANTTAAGTAGATGAGTGAAGAGTGAACGTGCTGCTTCCCTTTCCC | 180 | |
| OY | 181 TCCTGGCTACTGTCTCTTTGGACAACAATTCGCTGACTGTGACGTGACTTTAACG | 240 | |
| Db | 181 TCCTGGCTACTGTCTCTTTGGACAACAATTCGCTGACTGTGACGTGACTTTAACG | 240 | |
| OY | 241 AACGCCACATTCATTCATGAAGTTTCATACCCTATGACTGTCCAAGTCTGCTGTTCC | 300 | |

| | | | |
|----|------|---|------|
| Db | 241 | AACGCCACATCTCCATTTGAAGTTTATACCTATGCTGTCCACTGCTGCTTTCTCCC | 300 |
| Qy | 301 | CTTCAGACATATTTTGGCATCTGTTTCAACAGATGTCACACTGTCCTATGGAATACTG | 360 |
| Db | 301 | CTTCAGACATATTTTGGCATCTGTTTCAACAGATGTCACACTGTCCTATGGAATACTG | 360 |
| Qy | 361 | AAATGGAACAGTACTGTGGCAGTGAATGGAACAGCTTAATGGCAGCCCTGTAGAGGTTGGC | 420 |
| Db | 361 | AAATGGAACAGTACTGTGGCAGTGAATGGAACAGCTTAATGGCAGCCCTGTAGAGGTTGGC | 420 |
| Qy | 421 | AGTTTCCCCAGACTCCACGTTGTGGCATCAGGGGACGCTGATGGAACGTGTGTTTGT | 480 |
| Db | 421 | AGTTTCCCCAGACTCCACGTTGTGGCATCAGGGGACGCTGATGGAACGTGTGTTTGT | 480 |
| Qy | 481 | GGAATGACACAGTCATACAAATTATATAGATGTGTGTTAAAGATGCTCCTTGGCGG | 540 |
| Db | 481 | GGAATGACACAGTCATACAAATTATATAGATGTGTGTTAAAGATGCTCCTTGGCGG | 540 |
| Qy | 541 | CATGTGCATTTTCTCTTAATGGAAGCTTCTTTGTCACTGGCTCCTCATGTGTGATTTAA | 600 |
| Db | 541 | CATGTGCATTTTCTCTTAATGGAAGCTTCTTTGTCACTGGCTCCTCATGTGTGATTTAA | 600 |
| Qy | 601 | CAGTGTGGATGATTAATAATGAGTGTGATGTAGTGAATAAGCATTGATCTTGGAATTA | 660 |
| Db | 601 | CAGTGTGGATGATTAATAATGAGTGTGATGTAGTGAATAAGCATTGATCTTGGAATTA | 660 |
| Qy | 661 | CTTGCTCGATTTTCTTCAACAGCAGTTTCTGATGAGAAACAAGTCTTCAGTTTTC | 720 |
| Db | 661 | CTTGCTCGATTTTCTTCAACAGCAGTTTCTGATGAGAAACAAGTCTTCAGTTTTC | 720 |
| Qy | 721 | GACTGSCATCATGTGTCAAGATTTGCCAAGTCMAAATTTGATGTTCTTTTACCATA | 780 |
| Db | 721 | GACTGSCATCATGTGTCAAGATTTGCCAAGTCMAAATTTGATGTTCTTTTACCATA | 780 |
| Qy | 781 | TCCTT----- | 784 |
| Db | 781 | TCCTT----- | 784 |
| Qy | 785 | ----- | 784 |
| Db | 785 | ----- | 784 |
| Qy | 841 | CTTGCTGCTTTTCCCGATGGGAGATGCTAGTCTCAGGGTCAGTGAATAGCTGTCA | 900 |
| Db | 841 | CTTGCTGCTTTTCCCGATGGGAGATGCTAGTCTCAGGGTCAGTGAATAGCTGTCA | 900 |
| Qy | 901 | TAGATATGATTAATACTAGAAATATACTTCACACATTGACTCAGACACACAGTATG | 960 |
| Db | 901 | TAGATATGATTAATACTAGAAATATACTTCACACATTGACTCAGACACACAGTATG | 960 |
| Qy | 961 | TCACAACTGTGCTTTTGACCTAATACCCTTTTACTGTCTACGTGTTCAATGACAAA | 1020 |
| Db | 961 | TCACAACTGTGCTTTTGACCTAATACCCTTTTACTGTCTACGTGTTCAATGACAAA | 1020 |
| Qy | 1021 | CAGTGAACATCTGSCAATTTGACCTGGAACAATTGCCAAGCAAGGCGCACAGAACATC | 1080 |
| Db | 1021 | CAGTGAACATCTGSCAATTTGACCTGGAACAATTGCCAAGCAAGGCGCACAGAACATC | 1080 |
| Qy | 1081 | AGCTGAAGCAATTTACCGAAGATTGTCAGAGAGTCTGATCTCAACAATGCTTTGTGCAC | 1140 |
| Db | 1081 | AGCTGAAGCAATTTACCGAAGATTGTCAGAGAGTCTGATCTCAACAATGCTTTGTGCAC | 1140 |
| Qy | 1141 | AAAGTTTAAAGATCTGTGGTATTTTCAAGATGAATTAATGATGGAAGAACATCTGT | 1200 |
| Db | 1141 | AAAGTTTAAAGATCTGTGGTATTTTCAAGATGAATTAATGATGGAAGAACATCTGT | 1200 |
| Qy | 1201 | TGAATCTTAAAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTAGGACTGCGTA | 1260 |
| Db | 1201 | TGAATCTTAAAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTAGGACTGCGTA | 1260 |
| Qy | 1261 | GTAAGTGTGAGGAAATTTGAAGAGTCAAGACCAAGGTTAAATCCCTTTCTCAGGAA | 1320 |
| Db | 1261 | GTAAGTGTGAGGAAATTTGAAGAGTCAAGACCAAGGTTAAATCCCTTTCTCAGGAA | 1320 |
| Qy | 1321 | TTCTGTGATTTATATGTCCAAATACTAGAGAACTTATGAAAGATCCGCTATCGCAT | 1380 |
| Db | 1321 | TTCTGTGATTTATATGTCCAAATACTAGAGAACTTATGAAAGATCCGCTATCGCAT | 1380 |

Db 829 GAATTAATAATAAGTACAGTGAAGGCACTGCTCTCTGTTCTGCTTGTGCTTTT 888
Qy 785 ----- 784
Db 889 TCCCATGATGGGAGATGCTAGTCTCAGGGCTCAGTGATTAAGTCTGTCAATATATATGAT 948
Qy 785 ----- 784
Db 949 ACTAATACAGATATATCTTCAACATGTGACTGACACACAGGATATGTCAACATTTGT 1008
Qy 785 ----- 784
Db 1009 GCTTTGACCTAATACCTTTTACTGTCTACTGCTTCAATGACAAACAGTGAATC 1068
Qy 785 -----ACCAAGGGCCACAGAACTTCAGCTGAACCA 815
Db 1069 TGGCAATTGACCTGGAACAATTGGCCAGCAGGCGCACAGAACTCAGCTGAACCA 1128
Qy 816 TTTACCGAAGTTGGTCAGAGAGAGTGTCTCAACATGGCTTTTGTGCAAGATTTTAAA 875
Db 1129 TTTTACCGAAGTTGGTCAGAGAGAGTGTCTCAACATGGCTTTTGTGCAAGATTTTAAA 1188
Qy 876 GATCTTGTGGTATTTTCAAGATGAATPAACATTTGATGAAAAGAACTGTTGAATCTTACA 935
Db 1189 GATCTTGTGGTATTTTCAAGATGAATPAACATTTGATGAAAAGAACTGTTGAATCTTACA 1248
Qy 936 AAAGAAAGTCTGCTGATGATTTTGAATAATCTCTAGAGACTGCTGATGAAGTCTG 995
Db 1249 AAAGAAAGTCTGCTGATGATTTTGAATAATCTCTAGAGACTGCTGATGAAGTCTG 1308
Qy 996 AGGAATAATTGAAGCTCAGGACCAAGGTTAAATCCCTTTCTGAGGAATTCCTGATGA 1055
Db 1309 AGGAATAATTGAAGCTCAGGACCAAGGTTAAATCCCTTTCTGAGGAATTCCTGATGA 1368
Qy 1056 TTTTATGTCCAAATACATAGAGAACTTATGAAGATCCGGTCATCGATCGATGAGTGCAT 1115
Db 1369 TTTTATGTCCAAATACATAGAGAACTTATGAAGATCCGGTCATCGATCGATGAGTGCAT 1428
Qy 1116 TCATATGAAAAGGAAGCAATGGAATTTGATCAGCAAAAAGAAAGCTGACAGTCCCATG 1175
Db 1429 TCATATGAAAAGGAAGCAATGGAATTTGATCAGCAAAAAGAAAGCTGACAGTCCCATG 1488
Qy 1176 ACAAAATCTGTTCTCTCTGAGGGTACTTACACCAATATAGAGCTGGAATAATGCCATC 1235
Db 1489 ACAAAATCTGTTCTCTCTGAGGGTACTTACACCAATATAGAGCTGGAATAATGCCATC 1548
Qy 1236 AATGATGCTGAGAGACACCAAAAGTAAA 1266
Db 1549 AATGATGCTGAGAGACACCAAAAGTAAA 1579

RESULT 4
ADB47507
ID ADB47507 standard; cDNA; 1817 BP.
XX
AC ADB47507;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA upregulated in dendritic cells SEQ ID NO 207.
XX
KW 88; gene; human; dendritic cells; high throughput; cancer;
KW infectious disease; autoimmune disease; allergy;
KW graft versus host disease; vaccine enhancing; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003134283-A1.
XX
PD 17-JUL-2003.
XX
PF 03-OCT-2001; 2001US-00971392.
XX

PR 03-OCT-2000; 2000US-0237652P.
XX
XX (PETER) PETERSON D P.
PA (PEARL) PEARSON C I.
PA (COCK) COCKS B G.
XX
PI Peterson DP, Pearson CI, Cocks BG;
XX
DR WPI; 2003-662509/62.
XX
XX New combination comprises cDNAs that are differentially expressed in
PT dendritic cells useful for preparing a composition for diagnosing or
PT treating cancer, infectious disease, autoimmunity, allergy or graft
PT versus host disease.
XX
PS Claim 1; SEQ ID NO 207; 28pp; English.
XX
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in dendritic cells (DC). Also included is a high
CC throughput method for detecting differential expression of one or more
CC cDNAs in a sample containing nucleic acids. The combination is useful for
CC preparing a composition for diagnosing, treating and monitoring the
CC treatment of cancer, infectious disease, autoimmunity, allergy or graft
CC versus host disease, or for enhancing a vaccine. The present sequence
CC represents a human cDNA upregulated in dendritic cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docid=20030134283.
XX
SQ Sequence 1817 BP; 519 A; 373 C; 383 G; 542 T; 0 U; 0 Other;
Query Match 75.8%; Score 964.2; DB 10; Length 1817;
Best Local Similarity 81.8%; Pred. No. 5.2e-284;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;
Qy 12 TCACCTGCGGCGACGAGACCCGACCCCGTGGGACCTTGAAGCGGATCCCGGCG 71
Db 49 TCACCTGCGGCGACGAGACCCGACCCCGTGGGACCTTGAAGCGGATCCCGGCG 108
Qy 72 CCCCCGCTCTGCGAGGCTGTTTCTTCAATAAAGAAAGATGTAAGTCTGACACA 131
Db 109 CCCCCGCTCTGCGAGGCTGTTTCTTCAATAAAGAAAGATGTAAGTCTGACACA 168
Qy 132 TTAGCTGATGATGCTGAGAGATGCTCAACTGCTGCGCTTCTCTTCCCTTGGCTACT 191
Db 169 TTAGCTGATGATGCTGAGAGATGCTCAACTGCTGCGCTTCTCTTCTTGGCTACT 228
Qy 192 TGCTCTTGGACAAACAATTCGCTGTACTCGTTACGTGACTTTTACGTGACCTGCAAT 251
Db 229 TGCTCTTGGACAAACAATTCGCTGTACTCGTTACGTGACTTTTACGTGACCTGCAAT 288
Qy 252 TCTCATGTAAGATTTCATACCTATGCTGCTGCTGCTGTTTCTCCCTTCAGACAT 311
Db 289 TCTCATGTAAGATTTCATACCTATGCTGCTGCTGCTGTTTCTCCCTTCAGACAT 348
Qy 312 ATTTTGGCATGCTTTTAAAGATGCTGACAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 371
Db 349 ATTTTGGCATGCTTTTAAAGATGCTGACAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 408
Qy 372 ATGCTGCAATGATGAGAACAGCTGATGAGACCTGCTGAGAGGTTTCCAGTTTCCCA 431
Db 409 ATGCTGCAATGATGAGAACAGCTGATGAGACCTGCTGAGAGGTTTCCAGTTTCCCA 468
Qy 432 GACTCCAGTGTGTTGGCATGAGGGGAGCTGATGAACTGTGTTTGTGGAATGCAAG 491
Db 469 GACTCCAGTGTGTTGGCATGAGGGGAGCTGATGAACTGTGTTTGTGGAATGCAAG 528
Qy 492 TCATACAAATTAATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 551
Db 529 TCATACAAATTAATTAATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 588
Qy 552 TCTCCTAATGAAGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611

[illegible]

| | | |
|----------------------------|--|---------------|
| DT | 22-OCT-2001 | (first entry) |
| XX | | |
| DE | | |
| XX | | |
| XX | Human polynucleotide SEQ ID NO 1079. | |
| KW | Human; nocrotropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokineic; thrombolytic; drug screening; arthritis; inflammation; leukaemia; ss. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO200153312-A1. | |
| XX | | |
| PD | 26-JUL-2001. | |
| XX | | |
| PF | 26-DEC-2000; 2000WO-US034263. | |
| XX | | |
| PR | 23-DEC-1999; 99US-00471275. | |
| PR | 21-JAN-2000; 2000US-00488725. | |
| PR | 25-APR-2000; 2000US-0052317. | |
| PR | 20-JUN-2000; 2000US-00598042. | |
| PR | 19-JUL-2000; 2000US-00620312. | |
| PR | 03-AUG-2000; 2000US-00653450. | |
| PR | 14-SEP-2000; 2000US-00662191. | |
| PR | 19-OCT-2000; 2000US-00693036. | |
| PR | 29-NOV-2000; 2000US-00727344. | |
| XX | | |
| PA | (HYSE-) HYSEQ INC. | |
| XX | | |
| PI | Tang Y.T., Liu C., Agundi V., Chen R., Ma Y., Qian X.B., Ren F., Wang D.; Wang Y., Wang Z., Wehrman T., Xu C., Xue A.J., Yang Y., Zhang J., Zhao Q.A.; Zhou P., Goodrich R., Drmanac R.T. | |
| PI | | |
| DR | WP1; 2001-442253/47. | |
| DR | P-PSDBJ; AAM39720. | |
| XX | | |
| PT | Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries. | |
| XX | | |
| PS | Claim 1; SEQ ID NO 1079; 10078pp; English. | |
| XX | | |
| CC | The invention relates to human nucleic acids (AA157798-AA161369) and the encoded polypeptides (AAM38647-AA442213) with nocrotropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and centralised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S. disorders. Note: The sequence data for this patent did not form part of the printed specification | |
| CC | | |
| XX | | |
| SQ | Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other; | |
| XX | | |
| Query Match | 75.8%; Score 964.2; DB 4; Length 1844; | |
| Best Local Similarity | 81.8%; Pred. No. 5.2e-284; | |
| Matches 1252; Conservative | 0; Mismatches 3; Indels 276; Gaps 11; | |
| 0Y | 12 TCACCTGGCGGACGTGACCCGACCGCCCGTGGGACCTTGAAGCGGATCCGCGCG 71 | |
| DB | 65 TCACCTGGCGGACGTGACCCGACCGCCCGTGGGACCTTGAAGCGGATCCGCGCG 124 | |
| 0Y | 72 CCCCCGCTCCGCGAGCGCGCTTTTCTTCAATAAAGAAACGCGAAACGATTCACACA 131 | |
| DB | 125 CCCCCGCTCCGCGAGCGCGTGTTCCTTCAATAAAGAAACGCGGAACTGATTCACACA 184 | |

| | | | |
|----|------|--|------|
| OY | 132 | TTAGCTGACATCGTGTACATGTCACTGCTGTGGCTTCCTCTTCCCTTGGCTACT | 191 |
| Db | 185 | TTAGCTGACATCGTGTACATGTCACTGCTGTGGCTTCCTCTTCCCTTGGCTACT | 244 |
| OY | 192 | TGCTCTTGGACAAAACAATTGGCTGTACTGTGTACGTGACTTTCAGAACCTCAAT | 251 |
| Db | 245 | TGCTCTTGGACAAAACAATTGGCTGTACTGTGTACGTGACTTTCAGAACCTCAAT | 304 |
| OY | 252 | TCTCATTGAACTTTCATACCTATGCTGTTCACATGCTGTGCTTCTTCCCTTCAGACAT | 311 |
| Db | 305 | TCTCATTGAACTTTCATACCTATGCTGTTCACATGCTGTGCTTCTTCCCTTCAGACAT | 364 |
| OY | 312 | ATTTTGGCATCGNGTTCCAAACAATGGTAAACAACCTGTCCCTATGAAATPCTGAAAATGACAG | 371 |
| Db | 365 | ATTTTGGCATCGNGTTCCAAACAATGGTAAACAACCTGTCCCTATGAAATPCTGAAAATGACAG | 424 |
| OY | 372 | ATGCTGGCAGTATGAGAACAGCCTTAGTGGCAGCCTGTGAGAGGTTTGGCAGTTTCCCA | 431 |
| Db | 425 | ATGCTGGCAGTATGAGAACAGCCTTAGTGGCAGCCTGTGAGAGGTTTGGCAGTTTCCCA | 484 |
| OY | 432 | GACTCCACGTGTTTGGCATCAGGGGCACTGATGAACTGTGTGTTTGTGAAATGCACAG | 491 |
| Db | 485 | GACTCCACGTGTTTGGCATCAGGGGCACTGATGAACTGTGTGTTTGTGAAATGCACAG | 544 |
| OY | 492 | TCATCAAAATTATATGATGTGTACTGTATTAAGATGCTCTTGGGGGATGTGCATTT | 551 |
| Db | 545 | TCATCAAAATTATATGATGTGTACTGTATTAAGATGCTCTTGGGGGATGTGCATTT | 604 |
| OY | 552 | TCTCCTTAATGAAAGCTTCTTGTGCACTGTGCTCTCTCATGTGTGATTTTAAACAGTGGGAT | 611 |
| Db | 605 | TCTCCTTAATGAAAGCTTCTTGTGCACTGTGCTCTCTCATGTGTGATTTTAAACAGTGGGAT | 664 |
| OY | 612 | GATAAAATGAGGTGTCTGCATAGTGAATAAGCAATGATCTTGGAAATPACCTGTGGAT | 671 |
| Db | 665 | GATAAAATGAGGTGTCTGCATAGTGAATAAGCAATGATCTTGGAAATPACCTGTGGAT | 724 |
| OY | 672 | TTTTTCTTCAACCCAGTTTCTGATGAGAAACAAGGCTTCAGTTTGTGACATGGCATCA | 731 |
| Db | 725 | TTTTTCTTCAACCCAGTTTCTGATGAGAAACAAGGCTTCAGTTTGTGACATGGCATCA | 784 |
| OY | 732 | TGTGTGTCAGAGTTGCAAGTCAAAATTTGGATTGTTCTTTTACCCTATCTT----- | 784 |
| Db | 785 | TGTGTGTCAGAGTTGCAAGTCAAAATTTGGATTGTTCTTTTACCCTATCTT----- | 844 |
| OY | 785 | ----- | 784 |
| Db | 845 | GAATTAATAATATAAAGTACACTGAGTGGGCACTGTGCTCTGTTCGGCTTGCTTTT | 904 |
| OY | 785 | ----- | 784 |
| Db | 905 | TCCCATGATGGGCAGATGCTAGTCTCAGGGTCAGTGGAATTAAGTGTGATAGTATATGAT | 964 |
| OY | 785 | ----- | 784 |
| Db | 965 | ACTAACTACTGAGAATATACTTCAACACTGACTCAGACACAACAGTATGTCAACAATTGT | 1021 |
| OY | 785 | ----- | 784 |
| Db | 1025 | GCTTTTGCACCTAATACCTTTTACTTGTCTACTGTGTCAATGAGCAAAAACAGTGAACATC | 1081 |
| OY | 785 | -----AGCAAGGCGACAGAACATCAGCTGAAGCAA | 815 |
| Db | 1085 | TGGCAATTTGACCTGGAAAACATTTTGCAAAAGCAAGAGACACAAACATCAGCTGAAGCAA | 1144 |
| OY | 816 | TTTACCGAAGATTGGTCAGAGAGGTCGTCAACATGCTTGTGTGCAACAAGATTAAAA | 875 |
| Db | 1145 | TTTACCGAAGATTGGTCAGAGAGGTCGTCAACATGCTTGTGTGCAACAAGATTAAAA | 1200 |
| OY | 876 | GATCTTGTGTGATTTTCAAGAATGAATTAACATTTGAGAAAAAACAATGTTGAATCTTACA | 935 |
| Db | 1205 | GATCTTGTGTGATTTTCAAGAATGAATTAACATTTGAGAAAAAACAATGTTGAATCTTACA | 1266 |
| OY | 936 | AAAGAAAGTCTGGCTGATGATTTGAAAATTGAATCTCTAGAGACTCGGTATAAGTGTCTG | 995 |

| | | | |
|----------|---|---|------|
| Db | 1265 | AAAGAAAGTCTGGCTGATGATTTTGAAAATTGAACTCTAGACTGGCTAGTAAAGTCTG | 1324 |
| Oy | 996 | AGGAAAATTGAAGAGCTCAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCCTGATGAA | 1055 |
| Db | 1325 | AGGAAAATTGAAGAGCTCAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCCTGATGAA | 1384 |
| Oy | 1056 | TTTTATATGTCCAATTAACATAGAGAACTTATGAAGAATCCGGTCAATGCAATGAGTAT | 1115 |
| Db | 1385 | TTTATATATGTCCAATTAACATAGAGAACTTATGAAGAATCCGGTCAATGCAATGAGTAT | 1444 |
| Oy | 1116 | TCATATGAAAAAGAAAGCAATGAGAAATTTGGATCAGCAAAAAGAAACGTACAAATGCCATG | 1175 |
| Db | 1445 | TCATATGAAAAAGAAAGCAATGAGAAATTTGGATCAGCAAAAAGAAACGTACAAATGCCATG | 1504 |
| Oy | 1176 | ACAAATCTTGTCTTCTTCTTCTTCAAGCGGTACTTACACAAATAGAGACTGAAAATGGCCATC | 1235 |
| Db | 1505 | ACAAATCTTGTCTTCTTCTTCTTCAAGCGGTACTTACACAAATAGAGACTGAAAATGGCCATC | 1564 |
| Oy | 1236 | AATGATGCTGCGAGACACACCAAAAGTAAA | 1266 |
| Db | 1565 | AATGATGCTGCGAGACACACCAAAAGTAAA | 1595 |
| RESULT 6 | | | |
| ADQ99098 | | | |
| ID | ADQ99098 | standard; cDNA; 1844 BP. | |
| AC | ADQ99098; | | |
| XX | | | |
| DT | 23-SEP-2004 | (first entry) | |
| XX | | | |
| DE | DNA encoding human GPCR-like protein seqid 768. | | |
| XX | | | |
| KM | ophthalmological; immunomodulatory; cytoskeletal; antiatherosclerotic; | | |
| KM | antidiabetic; GPCR-like protein; ophthalmic disorder; | | |
| KM | neurological disorder; immunological disorder; nephritic disorder; | | |
| KM | hormonal dysfunction; cancer; atherosclerosis; diabetes; | | |
| OS | molecular weight marker; food supplement; human; ss. | | |
| XX | | | |
| XX | Homo sapiens. | | |
| PN | US6569662-B1. | | |
| PD | 27-MAY-2003. | | |
| XX | | | |
| PE | 19-JUL-2000; 2000US-00620312. | | |
| PR | 21-JAN-2000; 2000US-00488725. | | |
| PR | 25-APR-2000; 2000US-00552317. | | |
| XX | | | |
| PA | (HYSE-) HYSEQ INC. | | |
| XX | | | |
| PI | Tang YT, Zhou P, Dymnac RT; | | |
| DR | WPI; 2001-442255/47. | | |
| XX | | | |
| PT | New G-protein-coupled receptor-like polypeptides and polynucleotides. | | |
| PT | useful for treating diseases of ophthalmic, neurological, immunological | | |
| PT | and nephritic systems and hormonal dysfunction, cancer, atherosclerosis | | |
| PT | and diabetes. | | |
| XX | | | |
| PS | Example 2; SEQ ID NO 768; 92pp; English. | | |
| XX | | | |
| CC | The invention describes an isolated polynucleotide (I) comprising a fully | | |
| CC | defined (SI) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041, | | |
| CC | 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as | | |
| CC | given in the specification, its translated or protein coding portion, its | | |
| CC | extracellular portion or its active domain. The GPCR-like polypeptides | | |
| CC | and polynucleotides are useful for the treatment of diseases of | | |
| CC | ophthalmic, neurological, immunological and nephritic systems. They may | | |
| CC | also be used to treat hormonal dysfunction, cancer, atherosclerosis and | | |
| CC | diabetes. The antibodies are useful for detecting or quantitating the | | |

CC polypeptide in tissue. The polypeptides can also be used as molecular
CC weight markers and as a food supplement. This sequence represents a human
CC polynucleotide of the invention.

XX Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other;

Query Match 75.8%; Score 964.2; DB 5; Length 1844;
Best Local Similarity 81.8%; Pred. No. 5.2e-284;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

```
QY 12 TCACCTGCGCGGACGTCAGCCGACCCGCGGACCTTGAGGCGAGTCCCGCG 71
DB 65 TCACCTGCGCGGACGTCAGCCGACCCGCGGACCTTGAGGCGAGTCCCGCG 124
QY 72 CCCCCGCTCTGACGCGTGTCTTTCGAAATAAAGAACATGTGAACTGATTCA 131
DB 125 CCCCCGCTCTGACGCGTGTCTTTCGAAATAAAGAACATGTGAACTGATTCA 184
QY 132 TTAGCTGATCATTGTCACGATCTGTCAGTCTGCTGCTTCTTCTTCTTCTG 191
DB 185 TTAGCTGATCATTGTCACGATCTGTCAGTCTGCTGCTTCTTCTTCTTCTG 244
QY 192 TGCTCCTTGACAAAACAAATTCGCTGTACTGCTTACGATTTAGTGAATG 251
DB 245 TGCTCCTTGACAAAACAAATTCGCTGTACTGCTTACGATTTAGTGAATG 304
QY 252 TCTCATTGAAGTTTCAATACCTATGCTGCTGCTGCTTCTTCTTCTTCTG 311
DB 305 TCTCATTGAAGTTTCAATACCTATGCTGCTGCTGCTTCTTCTTCTTCTG 364
QY 312 ATTTTGGCATCGCTTCAACAGATGTCACATGCTGCTTACGAAATGTAAT 371
DB 365 ATTTTGGCATCGCTTCAACAGATGTCACATGCTGCTTACGAAATGTAAT 424
QY 372 ATGCTGGCAGTGTGAAACAGCCTAGTGGAGCCCTGAGGGGTTTCCGCA 431
DB 425 ATGCTGGCAGTGTGAAACAGCCTAGTGGAGCCCTGAGGGGTTTCCGCA 484
QY 432 GACTCAGCGTGTGAGCATGAGGGGACCTGATGAACTGTGTTTGTGAA 491
DB 485 GACTCAGCGTGTGAGCATGAGGGGACCTGATGAACTGTGTTTGTGAA 544
QY 492 TCATTAACAAATTAATGATGTGTGTAAGTGTGTAAGTGTGTAAGTGTG 551
DB 545 TCATTAACAAATTAATGATGTGTGTAAGTGTGTAAGTGTGTAAGTGTG 604
QY 552 TCTCCTAATGAGAGCTTCTTGTGCACTGCTCCTCATGTGTGTAATTA 611
DB 605 TCTCCTAATGAGAGCTTCTTGTGCACTGCTCCTCATGTGTGTAATTA 664
QY 612 GATTAATGAGAGTGTGCTGATGTAAGAAACATGATCTTGAATTAAC 671
DB 665 GATTAATGAGAGTGTGCTGATGTAAGAAACATGATCTTGAATTAAC 724
QY 672 TTTTCTTCAAGCAGTTTCTGATGGAACAAAGTCTTCAATTTTCACT 731
DB 725 TTTTCTTCAAGCAGTTTCTGATGGAACAAAGTCTTCAATTTTCACT 784
QY 732 TGTGTCAGAGTTGCCAAGTCAAAATTTGATGTTTCTTTACCAATAC 784
DB 785 TGTGTCAGAGTTGCCAAGTCAAAATTTGATGTTTCTTTACCAATAC 844
QY 785 ----- 784
DB 845 GAATTAATAATATAAGTACATGAGTGGGACCTGCTCTGTTCTGCTTGT 904
QY 785 ----- 784
DB 905 TCCCATGATGGGAGATGCTAGTCTCAGGGTCAAGTGAATAGTCTGTA 964
QY 785 ----- 784
DB 965 ACTAATAGTGAATATATCTTCAACATTTGATGATGATGATGATGATG 1024
```

```
QY 785 ----- 784
DB 1025 GCTTTTGACCTAATACCTTTTACTGCTAGTTCATGACAAAACAGTGA 1084
QY 785 ----- 784
DB 1085 TGGAATTTGACCTGGAACACCTTTCAGACAGGAGACATCATGCTGA 1144
QY 816 TTTACCGAAGATTGTCAGAGAGTGTCTCAACATGCTTTGTCAGAA 875
DB 1145 TTTACCGAAGATTGTCAGAGAGTGTCTCAACATGCTTTGTCAGAA 1204
QY 876 GATCTTGTGTTGTTTCAAGATGAACTTGAATGAAAGAACTTGAAT 935
DB 1205 GATCTTGTGTTGTTTCAAGATGAACTTGAATGAAAGAACTTGAAT 1264
QY 936 AAGAAAGTCTGGCTGATGATTTGAATTTGAATCTGAGGACTGGGTGA 995
DB 1265 AAGAAAGTCTGGCTGATGATTTGAATTTGAATCTGAGGACTGGGTGA 1324
QY 996 AAGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTTCTTCA 1055
DB 1325 AAGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTTCTTCA 1384
QY 1056 TTTATATGTCATTAATCTGAGAACTTATGAAAGATCCGCTATG 1115
DB 1385 TTTATATGTCATTAATCTGAGAACTTATGAAAGATCCGCTATG 1444
QY 1116 TCATTAAGAAAGGAGACCAATGGAATTTGATCAGCAAAAGAAAGT 1175
DB 1445 TCATTAAGAAAGGAGACCAATGGAATTTGATCAGCAAAAGAAAGT 1504
QY 1176 ACAATCTTGTCTTCTTCTTCAAGCGTACTTACCAAAATGAG 1235
DB 1505 ACAATCTTGTCTTCTTCTTCAAGCGTACTTACCAAAATGAG 1564
QY 1236 AATGATGCTGAGACACCAAAAGTAA 1266
DB 1565 AATGATGCTGAGACACCAAAAGTAA 1595

RESULT 7
ADB48858
ID ADB48858 standard; cDNA; 1844 BP.
XX
AC ADB48858;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human cDNA SEQ ID NO 768.
XX
KW 88; cancer; neurodegenerative disease; human.
XX
OS Homo sapiens.
XX
PN US2003104529-A1.
XX
PD 05-JUN-2003.
XX
PF 04-JAN-2002; 2002US-00037270.
XX
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 19-JUL-2000; 2000US-00620312.
XX
PA (ZHOU/) ZHOU P.
PA (TANG/) TANG Y T.
PA (LIU/) LIU C.
PA (ASUNDI/) ASUNDI V.
PA (DRMANAC/) DRMANAC R T.
XX
PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;
```


XX WPI; 2003-678194/64.
 XX New polynucleotide, useful for treating diseases e.g., cancer or
 PT neurodegenerative diseases.
 XX
 XX Claim 1; SEQ ID NO 768; 99bp; English.
 XX
 CC The invention relates to a polynucleotide comprising a sequence given in
 CC the specification, or its mature protein-coding portion, or its
 CC complement. The polynucleotide is useful for treating diseases e.g.,
 CC cancer or neurodegenerative diseases and many others listed in the
 CC specification. The present sequence represents a novel human cDNA. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format directly from USPTO
 CC at seqdata.uspto.gov/sequence.html?docID=20030104529.
 XX
 SQ Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other;
 Query Match 75.8%; Score 964.2; DB 9; Length 1844;
 Best Local Similarity 81.8%; Pred. No. 5.2e-284;
 Matches 1253; Conservative 0; Mismatches 3; Indels 276; Gaps 1;
 QY 12 TCACTCGCGGCGACGTGACCGCCCGTGGGACCTTGAAGCGGATCCCGGCG 71
 DB 65 TCACCTCGGGGCGACGTGACCGCCCGTGGGACCTTGAAGCGGATCCCGGCG 124
 QY 72 CCCCCGCTCTGCGAGGCTTTTCTTCAATAAAGAACATGTAATTCACACA 131
 DB 125 CCCCCGCTCTGCGAGGCTTTTCTTCAATAAAGAACATGTAATTCACACA 184
 QY 132 TTACGTATCATGTTGAGATGATGACGTGCTGCTTCTTCCCTTGGCTACT 191
 DB 185 TTACGTATCATGTTGAGATGATGACGTGCTGCTTCTTCCCTTGGCTACT 244
 QY 192 TGCTCTTGGACAAACAAATTCGCTTACTGTTAGTGAATCTGACAT 251
 DB 245 TGCTCTTGGACAAACAAATTCGCTTACTGTTAGTGAATCTGACAT 304
 QY 252 TCTTCATGAAGTTTCACTATGCTGCTGCTGCTGCTTCTTCCCTTCCGACAT 311
 DB 305 TCTTCATGAAGTTTCACTATGCTGCTGCTGCTGCTTCTTCCCTTCCGACAT 364
 QY 312 ATTTGGCATGTTTCAACAGATGTTACCACTGCTCCATGGAATPACTGAAATGACAG 371
 DB 365 ATTTGGCATGTTTCAACAGATGTTACCACTGCTCCATGGAATPACTGAAATGACAG 424
 QY 372 ATGCTGCAATGATGAAACAGCTAGTGGCAGCCCTGTGAGGGTTTCCGATTTCCCA 431
 DB 425 ATGCTGCAATGATGAAACAGCTAGTGGCAGCCCTGTGAGGGTTTCCGATTTCCCA 484
 QY 432 GACTCCACGTGTTGGCATCAGGGGAGCTGATGAACTGTGTTTGGAAATGACAG 491
 DB 485 GACTCCACGTGTTGGCATCAGGGGAGCTGATGAACTGTGTTTGGAAATGACAG 544
 QY 492 TCATACAAATTAATTAATGATGTTAGTTAAATGCTTCTTGGGCGATGTCATTT 551
 DB 545 TCATACAAATTAATTAATGATGTTAGTTAAATGCTTCTTGGGCGATGTCATTT 604
 QY 552 TCTCTAATGAAAGCTTCTTGTGATGCTCTCATGTTGATTTTAAACAGTGGGAT 611
 DB 605 TCTCTAATGAAAGCTTCTTGTGATGCTCTCATGTTGATTTTAAACAGTGGGAT 664
 QY 612 GATTAATAAGAGTGTGATGTAAGAAAGACATGATCTTGAATTAATCTGCTGCAT 671
 DB 665 GATTAATAAGAGTGTGATGTAAGAAAGACATGATCTTGAATTAATCTGCTGCAT 724
 QY 672 TTTTCTTCAAGCAGTTTCTGATGAGAAACAAGTCTTCAAGTTTTCAGCTGGCATCA 731
 DB 725 TTTTCTTCAAGCAGTTTCTGATGAGAAACAAGTCTTCAAGTTTTCAGCTGGCATCA 784
 QY 732 TGTGTGAGATGTCAGCAAGTCAAAATTTGATGTTTCTTTTACCATATCTT----- 784

DB 785 TGTGTGAGATGTCAGCAAGTCAAAATTTGATGTTTCTTTTACCATATCTT 844
 QY 785 ----- 784
 DB 845 GAATTAATAATTAAGATGATGATGAGGAGTGTCTTCTTCTGCTTGTGCTTTT 904
 QY 785 ----- 784
 DB 905 TCCCATGATGGGAGATGCTAGTCTGAGGCTGATGATGATGATGATGAT 964
 QY 785 ----- 784
 DB 965 ACTAATATGAGATATATCTTACACATGATGATGATGATGATGATGATGAT 1024
 QY 785 ----- 784
 DB 1025 GCTTTTGACCTTAATACCTTTTACTGTTGATGATGATGATGATGATGATGAT 1084
 QY 785 -----AGCAAGCGCACAGAACATCAGCTGAAACAA 815
 DB 1085 TGGCAATTTGACCTGAAACATCTTGGCAAGCAAGGACACAGAACATCAGCTGAAACAA 1144
 QY 816 TTACCGAAGATTTGTCAGAGAGTGTCTCAACATGCTTGTGTCACAAAGATTAAAA 875
 DB 1145 TTACCGAAGATTTGTCAGAGAGTGTCTCAACATGCTTGTGTCACAAAGATTAAAA 1204
 QY 876 GATTTTGTGTGATTTTCAAGATGATTAACATGATGATGATGATGATGATGATGAT 935
 DB 1205 GATTTTGTGTGATTTTCAAGATGATTAACATGATGATGATGATGATGATGATGAT 1264
 QY 936 AAGAAAGTCTGCTGATGATTTGAAATTAATCTCTAGGACTGCTGATGATGATGAT 995
 DB 1265 AAGAAAGTCTGCTGATGATTTGAAATTAATCTCTAGGACTGCTGATGATGATGAT 1324
 QY 996 AGGAAATTTGAAGAGTCAAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCGTATGAA 1055
 DB 1325 AGGAAATTTGAAGAGTCAAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCGTATGAA 1384
 QY 1056 TTTTATGTCATTAATTAAGAAATTCGCTGATGATGATGATGATGATGATGATGAT 1115
 DB 1385 TTTTATGTCATTAATTAAGAAATTCGCTGATGATGATGATGATGATGATGATGAT 1444
 QY 1116 TCATATGAAAGAGCAATGAAATTTGATGACGAAAGAAAGCAATGATGATGATGAT 1175
 DB 1445 TCATATGAAAGAGCAATGAAATTTGATGACGAAAGAAAGCAATGATGATGATGAT 1504
 QY 1176 ACAATCTTGTCTTCTTCTTCAAGGCTTACACCAATGATGATGATGATGATGATGAT 1235
 DB 1505 ACAATCTTGTCTTCTTCTTCAAGGCTTACACCAATGATGATGATGATGATGATGAT 1564
 QY 1236 AATGATGCTGAGACACACCAAAAGTAAA 1266
 DB 1565 AATGATGCTGAGACACACCAAAAGTAAA 1595
 RESULT 8
 AB224707 ID AB224707 standard; cDNA; 1773 BP.
 XX
 XX AB224707;
 AC
 XX
 DT 07-APR-2003 (first entry)
 XX
 XX Human cell growth, differentiation and death protein CGSD-19 cDNA.
 DE
 XX CGSD-19; cell growth; cell differentiation; cell death; human;
 KW cytosolic; antiarteriosclerotic; hepatotropic; antiinflammatory;
 KW antiproliferative; antianemic; ophthalmologic; auditory; anticonvulsant;
 KW cerebroprotective; nootropic; neuroprotective; antiParkinsonian;
 KW neuroleptic; tranquilizer; immunosuppressive; anti-HIV; antiallergic;
 KW antiaesthetic; antihypertensive; antidiabetic; dermatologic; nephroprotective;
 KW antipneumatic; antitachycardic; antitumor; antineoplastic; antineoplastic;
 KW antibacterial; fungicide; antiparasitic; protozoocidal; antihelminthic;

|||||
Db 1160 AAAAGAAAGTCTGCTGATGATTGAAATGAAATCTAGACTGCGTAGTAAGTCT 1219
Qy 995 GAGGAAATGAAGAGCTCAGACCAAGGTAATCCCTTTCTCAGAAATCTCTGATGA 1054
Db 1220 GAGGAAATGAAGAGCTCAGACCAAGGTAATCCCTTTCTCAGAAATCTCTGATGA 1279
Qy 1055 ATTATATATGTCATTAATCTAGAGAACTATGAAAGATCCGCTCATCGCATCAGATGCTA 1114
Db 1280 ATTATATGTCATTAATCTAGAGAACTATGAAAGATCCGCTCATCGCATCAGATGCTA 1339
Qy 1115 TTCAATGAAAGAGAGCAATGAAATTTGGATCAGCAAAAAGAAAGCTACAGTCCCAT 1174
Db 1340 TTCAATGAAAGAGAGCAATGAAATTTGGATCAGCAAAAAGAAAGCTACAGTCCCAT 1399
Qy 1175 GACAAATCTTGTCTTCTCTTACAGGGCTACTACACCAATGAGCTCTGAAATGGCCAT 1234
Db 1400 GACAAATCTTGTCTTCTCTTACAGGGCTACTACACCAATGAGCTCTGAAATGGCCAT 1459
Qy 1235 CAATAGATGGCTGAGACACACCAAAAGTANA 1266
Db 1460 CAATAGATGGCTGAGACACACCAAAAGTANA 1491

RESULT 9
ADB63535
ID ADB63535 standard; cDNA, 1811 BP.
XX
AC ADB63535;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA encoding clone TEST120264530.
XX
KM Human; ss: gene; pharmaceutical; diagnostic; gene therapy;
KM tissue regeneration; cell regeneration; membrane protein;
KM signal transduction-related protein; transcription-related protein;
KM osteoporosis; neurological disease; cancer; tumour.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 150..1580
FT /*cag= a
FT /product= "Clone TEST120264530 protein"
XX
PN EP1308459-A2.
XX
PD 07-MAY-2003.
XX
PE 28-MAR-2002; 2002EP-00007401.
XX
PR 05-NOV-2001; 2001JP-00379298.
PR 25-JAN-2002; 2002JUS-00350978.
XX
PA (HELI-) HELIX RES INST.
PA (REAS-) RES ASSOC BIOTECHNOLOGY.
XX
PI Iogaki T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I;
PI Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuho Y;
XX
XX WPI: 2003-450961/43.
XX P-PSDB; ADB65505.
XX
XX New polynucleotides and polypeptides, useful for developing a diagnostic
PT marker or medicines for regulation of their expression and activity, or
PT as targets of gene therapy.
XX
PS Claim 1; Page: 222pp; English.
XX
CC The invention discloses a polynucleotide comprising a sequence selected
CC from 1970 fully defined nucleotide sequences which encode novel

CC polypeptides. Also claimed is a polypeptide encoded by the polynucleotide
CC or its partial peptide, an antibody binding to the polypeptide or peptide
CC of the polynucleotide, immunologically assaying the polypeptide or peptide
CC peptide of the polynucleotide by contacting the polypeptide or peptide
CC with the antibody of the encoded protein, and observing the binding
CC between the two, a transformant carrying the polynucleotide in an
CC expressible manner and an antisense polynucleotide. The oligonucleotide
CC is useful as a primer for synthesizing the polynucleotide, or as a probe
CC for detecting the polynucleotide. The polynucleotides and encoded
CC proteins are useful as pharmaceutical agents and many disease-related
CC genes may be included in them, for developing a diagnostic marker or
CC medicines for regulation of their expression and activity, or as targets
CC of gene therapy. The genes are involved in tissue and/or cell
CC regeneration. Membrane proteins, signal transduction-related proteins,
CC transcription-related proteins, disease-related proteins and genes
CC encoding them can be used as indicators for diseases (e.g. osteoporosis,
CC neurological diseases, cancer, tumours). The cDNA may be used to regulate
CC the activity or expression of the encoded protein to treat diseases. The
CC sequence presented is a cDNA of the invention. Note: Some of the sequence
CC data for this patent is not represented in the printed specification, but
CC is based on sequence information supplied by the European Patent Office.
XX
SQ Sequence 1811 BP; 520 A; 352 C; 399 G; 540 T; 0 U; 0 Other;
Query Match 70.0%; Score 890.2; DB 10; Length 1811;
Best Local Similarity 80.9%; Pred. No. 2.4e-261;
Matches 1178; Conservative 0; Mismatches 3; Indels 276; Gaps 1;
Qy 86 GGCTGTTTTCTTCAATAAAGAACATGTAAGCTGAACTGATTCACATTAAGCTGATCAG 145
Db 125 GGCTGTTTTCTTCAATAAAGAACATGTAAGCTGAACTGATTCACATTAAGCTGATCAG 184
Qy 146 TGAGAGTGAATGCTGCTGCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 205
Db 185 TGAGAGTGAATGCTGCTGCTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 244
Qy 206 AACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 265
Db 245 AACATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304
Qy 266 TCATACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 325
Db 305 TCATACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 364
Qy 326 TTCAACAGATGATACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 385
Db 365 TTCAACAGATGATACCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 424
Qy 386 GGAACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 445
Db 425 GGAACAGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 484
Qy 446 GGCATCAGGGGAGCTGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 505
Db 485 GGCATCAGGGGAGCTGATGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 544
Qy 506 TAGATGCTGATGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 565
Db 545 TAGATGCTGATGCTGATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
Qy 566 CTTCCTTGTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 625
Db 605 CTTCCTTGTACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 664
Qy 626 TCTGATAGTGAAGAAAGCATGATCTGGAATTAACCTGCTGCTGCTGCTGCTGCTG 685
Db 665 TCTGATAGTGAAGAAAGCATGATCTGGAATTAACCTGCTGCTGCTGCTGCTGCTG 724
Qy 686 AGTTTCTGATGGAACAGAGCTCTTCACTTTTTCATGCTGATCATGCTGATGATG 745
Db 725 AGTTTCTGATGGAACAGAGCTCTTCACTTTTTCATGCTGATCATGCTGATGATG 784
Qy 746 CCAAGTCAAAATTTGATGATGCTTCTTTTACCAATCTT----- 784

```
Db 785 CCAAGTCAAAATTTGATGTTCTTTACCCCATCTTAGGTTGAATTAATATAA 844
QY 785 ----- 784
Db 845 AAGTACACTGAGTGGGCACTGCTCCTGTTCTGCGCTTGCTTTTCCCATGATGGCA 904
QY 785 ----- 784
Db 905 GATGCTAGTCTCAGGGTCAAGTGATAGTCTGATAGTATATGATTAATCTAGAA 964
QY 785 ----- 784
Db 965 TATACTTCAACATGTGACTGACACACAGTATGTCACAACTTGCTTTGCACCTAA 1024
QY 785 ----- 784
Db 1025 TACCCCTTTTACTGCTGCTGTTCAATGACAAACAGTGAACATCTGGCAATTTGACCT 1084
QY 785 -----AGCAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTG 829
Db 1085 GGAACACTTTGGCAAGCAAGAGACACAGAACATCAGCTGAAGCAATTTACCGAAGATTG 1144
QY 830 GTCAGAGAGTGTCTCAACATGCTGTTGTCACAAAGATTTTAAAGATCTTGCTGAT 889
Db 1145 GTCAGAGAGAGTGTCTCAACATGCTGTTGTCACAAAGATTTTAAAGATCTTGCTGAT 1204
QY 890 TTTCAAGATGATTAACATTTGATGATGAAAGAAAGAACTTTGATCTTCAAAAGAAAGTCTGCG 949
Db 1205 TTTCAAGATGATTAACATTTGATGATGAAAGAAAGAACTTTGATCTTCAAAAGAAAGTCTGCG 1264
QY 950 TGATGATTTGAAATTTGAATCTCTAGAGACTGCGGATGAAAGTCTGAGAAATTTGAGA 1009
Db 1265 TGATGATTTGAAATTTGAATCTCTAGAGACTGCGGATGAAAGTCTGAGAAATTTGAGA 1324
QY 1010 GCTCAGAGACCAAGGTTAAATCCCTTCTTCAAGAAATCTGATGAATTTATGCTCAAT 1069
Db 1325 GCTCAGAGACCAAGGTTAAATCCCTTCTTCAAGAAATCTGATGAATTTATGCTCAAT 1384
QY 1070 AACTAGAGAACTTTGAAAGATCCCGATCCGATCGATGAGTGGCTATTCATATGAAAGAA 1129
Db 1385 AACTAGAGAACTTTGAAAGATCCCGATCCGATCGATGAGTGGCTATTCATATGAAAGAA 1444
QY 1130 AGCAATGAAATTTGATCAGCAAAAGAAAGAAAGTACAGTCCCATGCAATCTTGTTCT 1189
Db 1445 AGCAATGAAATTTGATCAGCAAAAGAAAGAAAGTACAGTCCCATGCAATCTTGTTCT 1504
QY 1190 TCCTTCAAGCGGTACTTACACCAATAGAGACTTGAAATGCGCATCAATGATGCTGGA 1249
Db 1505 TCCTTCAAGCGGTACTTACACCAATAGAGACTTGAAATGCGCATCAATGATGCTGGA 1564
QY 1250 GACACACCAAAAGTAAA 1266
Db 1565 GACACACCAAAAGTAAA 1581

RESULT 10
ID ABV29028 standard; cDNA; 1996 BP.
XX AC ABV29028;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 29019.
XX KM Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KM pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX
```

```
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 6145-6146; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 1996 BP; 615 A; 378 C; 447 G; 543 T; 0 U; 13 Other;
XX
XX Query Match 69.1%; Score 878.8; DB 5; Length 1996;
XX Best Local Similarity 80.7%; Pred. No. 7.9e-258;
XX Matches 1166; Conservative 0; Mismatches 2; Indels 276; Gaps 1;
XX
XX 99 CAAATTAAGAACATGTGTGAAGTATTCACACATTAAGTATGATGATGAGATGTCAC 158
Db 4 CAAATTAAGAACATGTGTGAAGTATTCACACATTAAGTATGATGATGAGATGTCAC 63
QY 159 TGCTGTGCTTCTCTTTTCCCTTGGCTACTTGCTCTCTTGGACAAACATTTGCGCTG 218
Db 64 TGCTGTGCTTCTCTTTTCCCTTGGCTACTTGCTCTCTTGGACAAACATTTGCGCTG 123
QY 219 TACTGTTAGTGACTTACTGAACTGCAACATTCATGTTGCAATGTTTACTACTTGTCT 278
Db 124 TACTGTTAGTGACTTACTGAACTGCAACATTCATGTTGCAATGTTTACTACTTGTCT 183
QY 279 GTCCACTGTGCTGTTTCTCCCTTCAAGACATATTTGGACATGCTGTTCAACAGATGCT 338
Db 184 GTCCACTGTGCTGTTTCTCCCTTCAAGACATATTTGGACATGCTGTTCAACAGATGCT 243
QY 339 ACCACTGTCTTATGAAATCTGAAATTAAGACAGATGCTGCAATGTAACAGCTTACT 398
Db 244 ACCACTGTCTTATGAAATCTGAAATTAAGACAGATGCTGCAATGTAACAGCTTACT 303
QY 399 GGCAGCCCTGTAGGGTTTGGCAAGTTTCCCAAGATCCACAGGTTTGGCATAGGGGCA 458
Db 304 GGCAGCCCTGTAGGGTTTGGCAAGTTTCCCAAGATCCACAGGTTTGGCATAGGGGCA 363
QY 459 GCTGATGAACTGTGTTTGTGGAATGACAGTCAATACAAATTAATATGATGCTGATG 518
Db 364 GCTGATGAACTGTGTTTGTGGAATGACAGTCAATACAAATTAATATGATGCTGATG 423
QY 519 GTTAAAGATGCTCTTGTGCGGAGTGTGCAATTTTCTCTTAATGAAAGCTTCTTGTCACT 578
Db 424 GTTAAAGATGCTCTTGTGCGGAGTGTGCAATTTTCTCTTAATGAAAGCTTCTTGTCACT 483
```

| | | | |
|----|------|--|------|
| QY | 579 | GGCTCCATGCGGTGATTTAAACAATGGGATGATAAAGAGGTGTCGATAGTGA | 638 |
| Db | 484 | GGCTCCATGCGGTGATTTAAACAATGGGATGATAAAGAGGTGTCGATAGTGA | 543 |
| QY | 639 | AAAGCACAATGATCTTGGAAATTAACCTGCTGCGATTTTTCTTCACAGCCAGTTCTGATGGA | 698 |
| Db | 544 | AAAGCACAATGATCTTGGAAATTAACCTGCTGCGATTTTTCTTCACAGCCAGTTCTGATGGA | 603 |
| QY | 699 | GAACAAGGTCCTTCAGTTTTTTTCGACTGGGCATATGCGTCAGAGATTGCCAAGTCAAAATT | 758 |
| Db | 604 | GAACAAGGTCCTTCAGTTTTTTTCGACTGGGCATATGCGTCAGAGATTGCCAAGTCAAAATT | 663 |
| QY | 759 | TGATATGTTCTTCTTAAACCATTCTT | 784 |
| Db | 664 | TGATATGTTCTTCTTAAACCATTCTTAAAGTTAAATATAAAGTACACTGAGT | 723 |
| QY | 785 | ----- | 784 |
| Db | 724 | GGGCACTGCTGCTCTGTTCTGGCTTGCTTTTTCCCATGATGGGCAAGTGTAGTCTCA | 783 |
| QY | 785 | ----- | 784 |
| Db | 784 | GGGTCAATGATTAAGTCTGTCAATAGTATATGATTAATACTGAGAAATATATTCAACA | 843 |
| QY | 785 | ----- | 784 |
| Db | 844 | TTGACTACGACACACAGGATATGTCAACAATTGTGCTTTTGCACCTAATAACCTTTTACTT | 903 |
| QY | 785 | ----- | 784 |
| Db | 904 | GCTACTGCTCAATGACAAAAAGTGAACATCTGGCAATTTGACCTGGAAACCTTTGTC | 963 |
| QY | 785 | --ACCAAGGCCACAGAACATCACTGTAAGCAATTTACCGAAGATTGCTCAGAGAGGTC | 842 |
| Db | 964 | CAACCAAGGCCACAGAACATCACTGTAAGCAATTTACCGAAGATTGCTCAGAGAGAGAT | 1023 |
| QY | 843 | GTCTCAACATGGCTTTGTGCAACAAGATTAAAGATCTTGTGGTATTTTCAAGATGAT | 902 |
| Db | 1024 | GTCTCAACATGGCTTTGTGCAACAAGATTAAAGATCTTGTGGTATTTTCAAGATGAT | 1083 |
| QY | 903 | AACATGTGTAAGAAAGAACTGTGAACTTACAAAAGAAAGTCTGGCTGATGATTTGAA | 962 |
| Db | 1084 | AACATGTGTAAGAAAGAACTGTGAACTTACAAAAGAAAGTCTGGCTGATGATTTGAA | 1143 |
| QY | 963 | ATTGAATCTCAGGACCTGGCTAGTAAATGCTGAGGAAATTTGAAGAGCTCAGAGCCAG | 1022 |
| Db | 1144 | ATTGAATCTCAGGACCTGGCTAGTAAATGCTGAGGAAATTTGAAGAGCTCAGAGCCAG | 1203 |
| QY | 1023 | GTTAAATCCCTTTCTTCAGGAATTCCTGATGAATTTATGTCCAAATACTAGAGAACTT | 1082 |
| Db | 1204 | GTTAAATCCCTTTCTTCAGGAATTCCTGATGAATTTATGTCCAAATACTAGAGAACTT | 1263 |
| QY | 1083 | ATGAAGAATCCGGTTCATGCAATCAGATGGCTATTTCAATGAAAAAGAAAGCAATGGAAT | 1142 |
| Db | 1264 | ATGAAGAATCCGGTTCATGCAATCAGATGGCTATTTCAATGAAAAAGAAAGCAATGGAAT | 1323 |
| QY | 1143 | TGATATGACAAAAAGAAAGCTTACAAAGTCCCATGACAAATCTTGTCTTCTTCAGCGGTA | 1202 |
| Db | 1324 | TGATATGACAAAAAGAAAGCTTACAAAGTCCCATGACAAATCTTGTCTTCTTCAGCGGTA | 1383 |
| QY | 1203 | CTTACACCAATATAGGACTCTGAAAAATGGCCATCATATAGATGGCTGAGACACACCAAAAG | 1262 |
| Db | 1384 | CTTACACCAATATAGGACTCTGAAAAATGGCCATCATATAGATGGCTGAGACACACCAAAAG | 1443 |
| QY | 1263 | TAAA 1266 | |
| Db | 1444 | TAAA 1447 | |

| | | |
|----|---|-----|
| XX | ABV23190; | |
| AC | | |
| XX | 16-SEP-2002 (first entry) | |
| DT | | |
| XX | Human prostate expression marker CDNA 23181. | |
| DE | | |
| XX | Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker; | |
| KW | pharmacogenomic marker; gene; 88. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO200160860-A2. | |
| XX | | |
| PD | 23-AUG-2001. | |
| XX | | |
| PF | 20-FEB-2001; 2001WO-US005171. | |
| XX | | |
| PR | 17-FEB-2000; 2000US-0183319P. | |
| PR | 16-MAR-2000; 2000US-0189862P. | |
| PR | 25-MAY-2000; 2000US-0207454P. | |
| PR | 09-JUN-2000; 2000US-0211314P. | |
| PR | 18-JUL-2000; 2000US-0219007P. | |
| PR | 13-DEC-2000; 2000US-0255281P. | |
| XX | | |
| PA | (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC. | |
| XX | | |
| P1 | Schlegel R, Endege WO, Monahan JE; | |
| XX | | |
| DR | WPI; 2001-662795/76. | |
| XX | | |
| PT | Novel isolated nucleic acid molecule associated with cancerous state of | |
| PT | prostate cells and correlating with presence of prostate cancer, useful | |
| PT | for detecting presence of prostate cancer, stage of prostate cancer. | |
| XX | | |
| PS | Claim 1; Page 4166-4167; 11750pp; English. | |
| XX | | |
| CC | The invention relates to an isolated nucleic acid molecule (I) comprising | |
| CC | a nucleotide sequence given in Tables 1-9 (ABV000010-ABV62213) of the | |
| CC | specification or its complement. (I) is useful for: (a) assessing whether | |
| CC | a patient is afflicted with prostate cancer; (b) monitoring the | |
| CC | progression of prostate cancer in a patient; (c) assessing the efficacy | |
| CC | of a test compound to inhibit prostate cancer in a patient; (d) assessing | |
| CC | the efficacy of a therapy for inhibiting prostate cancer in a patient; | |
| CC | (e) selecting a composition for inhibiting prostate cancer in a patient; | |
| CC | (f) assessing the prostate cell carcinogenic potential of a compound; (g) | |
| CC | determining whether prostate cancer has metastasized in a patient; (h) | |
| CC | assessing the aggressiveness or incidence of prostate cancer in a patient | |
| CC | ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker | |
| XX | | |
| SQ | Sequence 1996 BP; 615 A; 378 C; 447 G; 543 T; 0 U; 13 Other; | |
| | | |
| | Query Match 69.1%; Score 878.8; DB 5; Length 1996; | |
| | Best Local Similarity 80.7%; Pred.No. 7.9e-258; | |
| | Matches 1166; Conservative 0; Mismatches 2; Indels 276; Gaps 1 | |
| OY | 99 CAAATTAAGAACATGTGTAACATGATTCACACATTAAGTCATGATGAGAGATGTCAAC | 158 |
| DB | 4 CAAATTAAGAACATGTGTAACATGATTCACACATTAAGTCATGATGAGAGATGTCAAC | 63 |
| OY | 159 TGTGTGCTTCTCTCTTTCCTCTTGGCTACTTGCTCTTGGACAAATAATTCGCTG | 218 |
| DB | 64 TGTGTGCTTCTCTCTTTCCTCTTGGCTACTTGCTCTTGGACAAATAATTCGCTG | 123 |
| OY | 219 TACTGTGTTAGTGACCTTTACTGAACTGCGACATTCCTCATTTGAAGTTTCATCTAGTCT | 278 |
| DB | 124 TACTGTGTTAGTGACCTTTACTGAACTGCGACATTCCTCATTTGAAGTTTCATCTAGTCT | 183 |
| OY | 279 GTTCACGTGCTGCTTTCTCCCTTGAAGACATATTTTGGCATGCTGTTCACAGATGCT | 338 |
| DB | 184 GTTCACGTGCTGCTTTCTCCCTTGAAGACATATTTTGGCATGCTGTTCACAGATGCT | 243 |
| OY | 339 ACCACTGCTCTATGAATACTGAATAATGACAGATGCTGGCATGATGTAACAGCTTAGT | 398 |

```

Db      244  ACCACGTCTATGAAATCTGAAAAATGACAGATGTGGCAGCTGATGGAACGCTAGT 303
QY      399  GGGAGCCCTGTGAGGGTTGGCCAGTTTCCCGAGCTCCAGCGTTTGGCATGAGGGCA 458
Db      304  GGGAGCCCTGTGAGGGTTGGCCAGTTTCCCGAGCTCCAGCGTTTGGCATGAGGGCA 363
QY      459  GCTGATGGAACCTGTGTTTGTGGAATGCACAGCTCATACAAATTATATAGATGTGAT 518
Db      364  GCTGATGGAACCTGTGTTTGTGGAATGCACAGCTCATACAAATTATATAGATGTGAT 423
QY      519  GTTAAAGATGGCTCTTGGCGGCAATGTCATTTTCTCTAATGGAAGCTTCTTGTCACT 578
Db      424  GTTAAAGATGGCTCTTGGCGGCAATGTCATTTTCTCTAATGGAAGCTTCTTGTCACT 483
QY      579  GGGTCCCATGATGATTTTAAACAGTGTGGATGATTAATGAGTGTGCAATGAGAA 638
Db      484  GGGTCCCATGATGATTTTAAACAGTGTGGATGATTAATGAGTGTGCAATGAGAA 543
QY      639  AAAGCATGATCTTGGAAATTACCTGCTGCAATTTTCTCAGAGCCAGTTTCTGATGA 698
Db      544  AAAGCATGATCTTGGAAATTACCTGCTGCAATTTTCTCAGAGCCAGTTTCTGATGA 603
QY      699  GAAACAAGTCTTCAAGTTTTCATGTCATCATGTGTCAGAGATTGCCAAGTCAAAAT 758
Db      604  GAAACAAGTCTTCAAGTTTTCATGTCATCATGTGTCAGAGATTGCCAAGTCAAAAT 663
QY      759  TGATGTTTCTTTTACCAATCTT----- 784
Db      664  TGATGTTTCTTTTACCAATCTT----- 723
QY      785  ----- 784
Db      724  GGGCATGTGCTCTCTGTTCTGGCTTGTCTTTTCCATGATGGCGAGATGCTAGTCA 783
QY      785  ----- 784
Db      784  GGGTCAGTGATTAAGTCTCATAGTATATGATCTAATATGAGAAATATATCTACACA 843
QY      785  ----- 784
Db      844  TTGACTCAGACACACAGTATGTCAACAATGTGCTTTGACCTAATACCTTTTACTT 903
QY      785  ----- 784
Db      904  GCTACTGTTCAATGACCAAAAACAGTAACATCTGGCAATTGACCTGGAACAATTGC 963
QY      785  --AGCAAGGGCGACAGAAATCAGCTGAGCAATTTACGAGAATTTGTCAGAGAGTC 842
Db      964  CAAAGCAAGGGCGACAGAAATCAGCTGAGCAATTTACGAGAATTTGTCAGAGAGAT 1023
QY      843  GTCTCAACATGCTTGTGCACAAGATTTTAAAGATCTTGTGTTTCAAGATCAAT 902
Db      1024  GTCTCAACATGCTTGTGCACAAGATTTTAAAGATCTTGTGTTTCAAGATCAAT 1083
QY      903  AACATGATGAAAAAGAACTGTGAATCTTACAAAAAGAAAGTGTGCTGATGATTTGAA 962
Db      1084  AACATGATGAAAAAGAACTGTGAATCTTACAAAAAGAAAGTGTGCTGATGATTTGAA 1143
QY      963  ATTGAATCTTAGAGCTGCGTAGTAAGATGCTGAGAAAAATTGAAGACTCAGAGCAAG 1022
Db      1144  ATTGAATCTTAGAGCTGCGTAGTAAGATGCTGAGAAAAATTGAAGACTCAGAGCAAG 1203
QY      1023  GTTAAATCCCTTTTTCAGAAATCTCTGATGAAATTTATGTCCAAATTAACATGAGAACT 1082
Db      1204  GTTAAATCCCTTTTTCAGAAATCTCTGATGAAATTTATGTCCAAATTAACATGAGAACT 1263
QY      1083  ATGAAATATCGGTCATCGCATGAGATGGCATTCATATGAAAGAAAGCAATGGAAT 1142
Db      1264  ATGAAATATCGGTCATCGCATGAGATGGCATTCATATGAAAGAAAGCAATGGAAT 1323
QY      1143  TGGATGCAAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTCTTCCACCGGTA 1202

```

```

Db      1324  TGGATCAGCAAAAAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTCTTACAGCGGTA 1383
QY      1203  CTTACACCAAAATAGGACTCTGAAAAATGGCCATCAATAGATGGCTGAGACACACCAAAAG 1262
Db      1384  CTTACACCAAAATAGGACTCTGAAAAATGGCCATCAATAGATGGCTGAGACACACCAAAAG 1443
QY      1263  TAAA 1266
Db      1444  TAAA 1447

RESULT 12
AAD45076
ID      AAD45076 standard; DNA; 1908 BP.
XX
AC      AAD45076;
XX
DT      27-DEC-2002 (first entry)
XX
DE      Human RET16.3 splice variant DNA.
XX
KW      Human; RET16; intracellular signal; inflammation-related disease; asthma;
KW      rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective;
KW      transplant rejection; chronic obstructive pulmonary disease; TNF-alpha;
KW      inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
KW      acute respiratory distress syndrome; cardiac; ulcerative colitis;
KW      autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm;
KW      cellular migration disorder; cell proliferation disorder; calcification;
KW      hyperinsulinaemia; diabetes type 2; systemic lupus erythematosus; tumour;
KW      cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
KW      thalassemia; vasculitic; gene; ds.
XX
OS      Homo sapiens.
XX
FH      Key      Location/Qualifiers
FT      CDS      136..1644
FT      /tag= a
FT      /product= "Human RET16.3 splice variant protein"
XX
PN      WO200266494-A2.
XX
PD      29-AUG-2002.
XX
PF      15-FEB-2002; 2002WO-US005162.
XX
PR      16-FEB-2001; 2001US-0269366P.
PR      29-MAY-2001; 2001US-0294181P.
XX
PA      (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI      Todderud CG, Finger JN, Rillema J;
XX
DR      WPI; 2002-682760/73.
XX
DR      P-PSDB; AAE28168.
XX
PT      New human, mouse or rat RET16 genes and proteins, involved in
PT      intracellular signaling cascade, useful for in gene therapy, particularly
PT      for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
PT      tumors or neoplasms.
XX
PS      Claim 1; Page 168-169; 175p; English.
XX
CC      The invention relates to human, mouse or rat RET16 genes and proteins,
CC      involved in intracellular signaling cascade. The RET16 protein or
CC      polynucleotide is useful for treating an inflammation-related disease or
CC      disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis,
CC      asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or
CC      tissue transplants, chronic obstructive pulmonary disease, inflammatory
CC      bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory
CC      distress syndrome, systemic lupus erythematosus, autoimmune disease,
CC      cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
CC      disease or disorder also includes disorders associated with aberrant
CC      activation of the TNF-alpha pathway, disorders associated with aberrant

```


XX 26-DEC-2000; 2000MO-US034263.
PF
XX
PR 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
PI Tang YT, Lau Z, Ahundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao Q;
PI Zhou P, Goodrich R, Drmanac RT;
XX
DR MPI; 2001-442253/47.
XX P-PSDB; AAM41506.
XX
PT Novel nucleic acids and polypeptides, useful for treating disorders such
as central nervous system injuries.
XX
XX Claim 1; SEQ ID NO 4651; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
encoded polypeptides (AAM38642-AAM42213) with nootropic,
immunosuppressant and cytoskeletal activity. The polynucleotides are useful
in gene therapy. A composition containing a polypeptide or polynucleotide
of the invention may be used to treat diseases of the peripheral nervous
system, such as peripheral nervous injuries, peripheral neuropathy and
localised neuropathies and central nervous system diseases, such as
Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
inhibition of the activities such as: Immune system suppression,
CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
and thrombolytic activity, cancer diagnosis and therapy, drug screening
and assays for receptor activity, arthritis and inflammation, leukaemia and
CC C.N.S disorders. Note: The sequence data for this patent did not form
part of the printed specification
XX
XX Sequence 1826 BP; 535 A; 382 C; 380 G; 529 T; 0 U; 0 Other;
SQ
Query Match 68.2%; Score 867.8; DB 4; Length 1826;
Best Local Similarity 79.6%; Pred. No. 1.8e-254;
Matches 1232; Conservative 0; Mismatches 27; Indels 288; Gaps 6;
QY 12 TCACCTGCGCGGACGTCACCGCCGCGGACCTTGAAAGCGGATCCGCGCG 71
DB 1800 TCACCTGCGCGGACGTCACCGCCGCGGACCTTGAAAGCGGATCCGCGCG 1741
QY 72 CCCCCGCTCTCGACGCTGTTTTCTTCAATAAAGACATGTTGAATCTACACA 131
DB 1740 CCCCCGCTCTCGACGCTGTTTTCTTCAATAAAGACATGTTGAATCTACACA 1681
QY 132 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 131
DB 1680 TTAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1621
QY 192 TGTCTCTTGACAAACAAATTCGCTGATCTGTTAGTGAATTTAGTGAATGCGCAT 251
DB 1620 TGTCTCTTGACAAACAAATTCGCTGATCTGTTAGTGAATTTAGTGAATGCGCAT 1561
QY 252 TCTTCATTGAAGTTTCACTACCTATGCTTCACCTGCTGCTGTTCTCCCTTCAGACAT 311
DB 1560 TCTTCATTGAAGTTTCACTACCTATGCTTCACCTGCTGCTGTTCTCCCTTCAGACAT 1501
QY 312 ATTTGGCATCGTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 371
DB 1500 ATTTGGCATCGTGTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1441

QY 372 ATGCTGGCAGTATGGAACAGCCCTAGTGGCAGCCCTGTGAGGGTTTGCCAGTTTCCCA 431
DB 1440 ATGCTGGCAGTATGGAACAGCCCTAGTGGCAGCCCTGTGAGGGTTTGCCAGTTTCCCA 1381
QY 432 GACTCCAGCTGTTTGGCATCAGGGGACGATGGAACGTGGTTTGTGGAAATGACAG 491
DB 1380 GACTCCAGCTGTTTGGCATCAGGGGACGATGGAACGTGGTTTGTGGAAATGACAG 1321
QY 492 TCATCAAAATTATATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 551
DB 1320 TCATCAAAATTATATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1261
QY 552 TCTCTTAATGAAGCTTCTTTGTCACTGCTCTCATGTGTGATTTAAACATGCGGAT 611
DB 1260 TCTCTTAATGAAGCTTCTTTGTCACTGCTCTCATGTGTGATTTAAACATGCGGAT 1201
QY 612 GATTAATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 671
DB 1200 GATTAATAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1141
QY 672 TTTCTTCAAGCAGCTTCTGATGAGAACAGGCTTCAGTTTTCGACTGCGATCA 731
DB 1140 TTTCTTCAAGCAGCTTCTGATGAGAACAGGCTTCAGTTTTCGACTGCGATCA 1081
QY 732 TGTGTCAGGATTTGCCAAGTCAAAATTTGGATTTTCTTTAACCATACTT----- 784
DB 1080 TGTGTCAGGATTTGCCAAGTCAAAATTTGGATTTTCTTTAACCATACTTAGGTTT 1021
QY 785 ----- 784
DB 1020 GAATTAATAATTAATAAGTACACTGATGAGGCACTGTCTCTGTTGCTGTGCTTTT 961
QY 785 ----- 784
DB 960 TCCCATGATGGGACATGCTAGTCTAGGGTCAATGATTAAGTCTCATAGTATATGAT 901
QY 785 ----- 784
DB 900 ACTAATAGTGAATATATCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 841
QY 785 ----- 784
DB 840 GCTTTTGACCTAATACCTTTTACTTGTCTAGTGTCAATGACAAAACATGAACATC 781
QY 785 -----AGCAAGCGCACAGAACATGATGAGCA 815
DB 780 TGGCAATTTGACCTGGAACACCTTTCGCAAGCAAGACACAGAAATCAGCTGAAGCA 721
QY 816 TTTACCGAAGATTGTCAGAGAGTCTCTCAACATGCTTTTGTGCACAAAGTTTAA 875
DB 720 TTTACCGAAGATTGTCAGAGAGTCTCTCAACATGCTTTTGTGCACAAAGTTTAA 661
QY 876 GATCTTGTGATTTTCAAGATGATTAACATTTGATGAAAGAACTGTTGAATCTTACA 935
DB 660 GATCTTGTGATTTTCAAGATGATTAACATTTGATGAAAGAACTGTTGAATCTTACA 601
QY 936 AAAGAAGTCTGCTGATGATTTGAATTTCTAGAGATGCTGATGATGATGATGATGATGATGATGAT 995
DB 600 AAAGAAGTCTGCTGATGATTTGAATTTCTAGAGATGCTGATGATGATGATGATGATGATGATGAT 541
QY 996 AGGAATTTGAAGCTCAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCCTGATGA 1055
DB 540 AGGAATTTGAAGCTCAGACCAAGTTAAATCCCTTTCTTCAAGAAATTCCTGATGA 481
QY 1056 TTTATATGTCATTAATCTAGAGACTTATGAAAGATCCGCTCATCGCATCAGTGGCTAT 1115
DB 480 TTTATATGTCATTAATCTAGAGACTTATGAAAGATCCGCTCATCGCATCAGTGGCTAT 421
QY 1116 TCATATGAAGAAGCAATGGAATTTGG--ATCAGCAAAAAGAA--GATACAGTCC 1172
DB 420 TCATATGAAGAAGCAATGGAATTTGGATTCAGCAAAAAGAA--GATACAGTCC 361
QY 1173 ATGACAAATCTTGTCTCTT-----CAGGCTAC--TTACACCAATAGGACTCTGA 1224

Db 1381 GTAAA 1385
RESULT 15
ADM19385 standard; cDNA; 1686 BP.
ID ADM19385
AC ADM19385;
XX
XX
DT 20-MAY-2004 (first entry)
DE Novel human channel/transporter gene #182.
XX
XX de; gene; immunosuppressive; antiarthritic; antirheumatic;
KM antiproliferative; cyostatic; cardiant; vasotropic; cerebroprotective;
KM nootropic; neuroprotective; antibacterial; vitucide; fungicide;
KM optalmological; gene therapy; channel/transporter protein;
KM rheumatoid arthritis; neoplaem; cardiac arrest; cerebrovascular disorder;
KM cerebral ischemia; angiogenesis; nervous system disorder;
KM Alzheimer's disease; ocular disorder; corneal infection; wound healing;
KM epithelial cell proliferation; skin aging; sunburn; transplantation;
KM chemotaxis; food additive.
XX
OS Homo sapiens.
XX
XX WO200154472-A2.
XX
XX PD 02-AUG-2001.
XX
XX PF 17-JAN-2001; 2001WO-US001307.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234474P.
PR 25-SEP-2000; 2000US-0234897P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.

[illegible]

```

Db      1251 AAGGAAAGCTGGCTGATGATTTGAAAATTGATGGCTATTCAATATGAAAAAGAACATG 1310
QY      1137 GAAATTTGATCAGCAAAAAAGAACGTACAGTCCCATGACAAATCTTGTCTTCTTCA 1196
Db      1311 GAAATTTGATCAGCAAAAAAGAACGTACAGTCCCATGACAAATCTTGTCTTCTTCA 1370
QY      1197 GCGGTACTTACCAAAATAGGACTCTGAAAAATGGCCATCAATAGATGGCTGAGACACAC 1256
Db      1371 GCGGTACTTACCAAAATAGGACTCTGAAAAATGGCCATCAATAGATGGCTGAGACACAC 1430
QY      1257 CAAAAGTAAA 1266
Db      1431 CAAAAGTAAA 1440

```

Search completed: February 5, 2005, 05:27:38
 Job time : 789 secs

transplant rejection; chronic obstructive pulmonary disease; TNF-alpha; inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer; inacute respiratory distress syndrome; cardiac; ulcerative colitis; autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm; cellular migration disorder; cell proliferation disorder; calcification; hyperinsulinemia; diabetes type 2; systemic lupus erythematosus; tumour; cardiovascular disease; Wegener's granulomatosis; atherosclerosis; thalassemia; vasotropic; open reading frame; ORF; gene; ds.

OS Homo sapiens.

XX Key Location/Qualifiers

XX CDS 3..1541

FT /*tag= a

FT /product= "Human RET16 protein"

FT /note= "CDS does not include start codon"

FT /partial

XX MO20266494-A2.

XX PD 29-AUG-2002.

XX PF 15-FEB-2002; 2002W0-US005162.

XX PR 16-FEB-2001; 2001US-0269366P.

XX PR 29-MAY-2001; 2001US-0294181P.

XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.

XX PI Todderud CG, Finger JN, Rillema J;

XX DR WPI: 2002-682760/73.

XX DR P-PSDB; AAE28164.

XX PT New human, mouse or rat RET16 genes and proteins, involved in intracellular signaling cascade, useful for in gene therapy, particularly for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers, tumours or neoplasms.

XX PS Claim 1; Page 150; 175pp; English.

XX CC The invention relates to human, mouse or rat RET16 genes and proteins, involved in intracellular signaling cascade. The RET16 protein or polynucleotide is useful for treating an inflammation-related disease or disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis, asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or tissue transplants, chronic obstructive pulmonary disease, inflammatory bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory distress syndrome, systemic lupus erythematosus, autoimmune disease, cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related disease or disorder also includes disorders associated with aberrant activation of the TNF-alpha pathway, disorders associated with aberrant cellular migration, proliferation, metastasis, juvenile idiopathic arthritis, haematogenous metastases of tumour cells, hyperinsulinemia, diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer, CC tumour progression, Wegener's granulomatosis, stem cell transplantation complications, ischaemia-reperfusion injury, thalassemia, acute lung CC injury, graft rejection, ischaemic heart, coronary artery calcification CC or allergic inflammation. RET16 DNA is used in gene therapy. The present CC sequence is human RET16 open reading frame (ORF) DNA

XX SQ Sequence 1553 BP; 428 A; 327 C; 348 G; 450 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.86e-185 Length: 1553

Score: 1984.00 Matches: 383

Percent Similarity: 80.46% Conservative: 0

Best Local Similarity: 80.46% Mismatches: 1

Query Match: 96.92% Indels: 92

DB: Gaps: 1

US-10-077-111-13 (1-384) x AAD45071 (1-1553)

Qy 1 MetValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValaenCysCysAlaPhe 20

Db 111 ATGTGTAAACTGATTCACATTAAGCTGATCAATGTCATGCTGCTGCTTC 170

Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40

Db 171 TCTTTTCCCTTGGCTGCTACTGCTCTTGGACAAAACAAATGGCCGTGACTGTTACGT 230

Qy 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys 60

Db 231 GACTTAACTGAACCTGCACATTCCTCAATTGAAGTTTCAATGCTGCTGCTGCTGC 290

Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80

Db 291 TGTTCCTCCCTTCAGGACATATTTTGGCATGCTGTTCAACAGATGGTACCAGTCTCTA 350

Qy 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValaMetGluGlnProSerGlySerProVal 100

Db 351 TGGAAATCTGAATAATGACAGATGCTGGCAGTATGGAAACAGCTGATGGCAGCCCTGTG 410

Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValaAlaAspGlyThr 120

Db 411 AGGATTGGCCAGTTTTCCTCCAGACTCCAGTGTGTCATCAGGAGCCTGATGAACT 470

Qy 121 ValValLeuTrpAsnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly 140

Db 471 GTGTTTGTGGAAATGACAGCTCATACAAATTATATGATGTGTAGTGTAAAGATGCG 530

Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160

Db 531 TCTTGGCGGCAATGTGATTTTCTCTTAATGAAAGCTTTCTTGTCACTGCTCCTCATGT 590

Qy 161 GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyValaAlaAsp 180

Db 591 GGTGATTAAACAGTGTGGAGTGAATAATAGGTGTCTGATGTGAAAAAGCACATGAT 650

Qy 181 LeuGlyLethrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200

Db 651 CTGGAAATACCTGCTGCGCATTTTCTTCACAGCCAGTTTCTGTGAGAAACAGTCTT 710

Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer 220

Db 711 CAGTTTGTGACCTGGCATATGTGGTCAGAGATTGCCAATCAAAANTTGGATTGTCT 770

Qy 221 PheThrHisIleLeu----- 225

Db 771 TTAAACCATATCTTAGCTTTGAATTAAATATAAAGTACACTGAGTGGCACTGTGCT 830

Qy 225 ----- 225

Db 831 CTTGTTCTGCTTGTGCTTTTCCCGTATGGCGAGATGCTAGTCTCAGGCTCAGTGGAT 890

Qy 225 ----- 225

Db 891 AAGTCGTGATAGTATATGATTAATTAATACGAGAATATTACTGCACACATTGACACGAC 950

Qy 225 ----- 225

Db 951 ACCAGTATGTCAAACTTGTGCTTTTGCACCTAATACCTTTTACTGCTACTGATGTTCA 1010

Qy 226 -----AlaArgArg 228

Db 1011 ATGACAAAACAGTGAACATCTGGCAATTGACCTGGAAACACTTGGCCAGCAAGGCGC 1070

Qy 229 ThrGlnHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValLysThrTrp 248

Db 1071 ACGAACAATACAGCTGAAGCAATTTACCGAAGATTGGTCAGAGAGAGATGCTCAACATGG 1130

Qy 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268

Db 1131 CTTTGTGACACAAATATTAAAGATCTTGTGTATTTTTCAGATGATTAACATTTGATGGA 1190

Qy 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288

Db 1191 AAGAACTGTGATATCTACAAAAGAGTGGCTGATGATTTGAAAATTGAAATCTCTA 1250
Qy 289 G1YleuAserSerlyValIeuArlgLyIleG1uGlueuArlgThrluValIysSerIeu 308
Db 1251 GGGCTGGCTGTAAAGTCTGAGGAAATGAAAGCTCAGGACCAAGTTAAATCCCTT 1310
Qy 309 SerSerG1YIleProApgIupheIleCyAProIleThrrAgIueuWetLyAAsPro 328
Db 1311 TCCTCAGAAATCTCGATGAAATTTATATGTCCAATACTAGAGAACTTATGAAAGATCCG 1370
Qy 329 ValIleAlaSerSerPglYTrSerTYrGluLyGluAlaMetGluAuantP1IleSerLy 348
Db 1371 GTCATGCGCATCAATGCTCTATTCAATGAAAGAACCAATGGAAATTGATCAGCAAA 1430
Qy 349 LyIleYArgrThrSerProMetCTrrAsnIleuValIeuProSerAlaValIeuthrProAsn 368
Db 1431 AAGAAAGCTGACAAAGTCCCAAGCAAAATCTGTTCTTCTTCAGCGGACTTACCAAAAT 1490
Qy 369 ArgThrluYleuMetAlaIleAsnArlgTrpLeuGluThrluIsgIuLy 384
Db 1491 AGGACTCTGAATAATGGCATCAATAGATGCTGGAGACACCAAAAG 1538
RESULT 3
AAD45070
ID AAD45070 standard; DNA; 1818 BP.
XX
AC AAD45070;
XX
DT 27-DEC-2002 (first entry)
XX
DE Human RET16 DNA.
XX
XX Human; RET16; intracellular signal; inflammation-related disease; asthma;
XX rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotection;
XX transaplar rejection; chronic obstructive pulmonary disease; TNF-alpha;
XX inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer;
XX inactive respiratory distress syndrome; cardiac; ulcerative colitis;
XX acute immune disease; cystic fibrosis; gene therapy; cytotoxic; neoplasm;
XX cellular migration disorder; cell proliferation disorder; calcification;
XX hyperinflammation; diabetes type 2; systemic lupus erythematosus; tumour;
XX cardiovascular disease; Wegener's granulomatosis; atherosclerosis;
XX thalassemia; vasotropic; gene; db.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS /tag= a
FT /product= "Human RET16 protein"
XX
XX MO200266494-A2.
XX
XX PD 29-AUG-2002.
XX
XX PF 15-FEB-2002; 2002MO-US005162.
XX
XX PR 16-FEB-2001; 2001US-0269366P.
XX PR 29-MAY-2001; 2001US-0294181P.
XX
XX PA (BRIM) BRISTOL-MYERS SQUIBB CO.
XX
XX PI Todderud CG, Finger JN, Rillema J;
XX
XX DR WPI; 2002-682760/73.
XX
XX DR P-PSDB; AAE28163.
XX
XX PT New human, mouse or rat RET16 genes and proteins, involved in
XX intracellular signaling cascade, useful for in gene therapy, particularly
XX for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
XX tumors or neoplasms.
XX
XX PS Claim 1; Page 147-148; 175pp; English.

XX
CC The invention relates to human, mouse or rat RET16 genes and proteins,
CC involved in intracellular signaling cascade. The RET16 protein or
CC polynucleotide is useful for treating an inflammation-related disease or
CC disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis,
CC asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or
CC tissue transplants, chronic obstructive pulmonary disease, inflammatory
CC bowel disease, Crohn's disease, ulcerative colitis, inactive respiratory
CC distress syndrome, systemic lupus erythematosus, autoimmune disease,
CC cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
CC disease or disorder also includes disorders associated with aberrant
CC activation of the TNF-alpha pathway, disorders associated with aberrant
CC cellular migration, proliferation, metastasis, juvenile idiopathic
CC arthritis, haematogenous metastases of tumour cells, hyperinflammation,
CC diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer,
CC tumour progression, Wegener's granulomatosis, stem cell transplantation
CC complications, ischaemia-reperfusion injury, thalassemia, acute lung
CC injury, graft rejection, ischaemic heart, coronary artery calcification
CC or allergic inflammation. RET16 DNA is used in gene therapy. The present
CC sequence is human RET16 DNA
XX

Sequence 1818 BP; 523 A; 373 C; 382 G; 540 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 3,58e-185 | Length: | 1818 |
| Score: | 1984.00 | Matches: | 383 |
| Percent Similarity: | 80.46% | Conservative: | 0 |
| Best Local Similarity: | 80.46% | Mismatches: | 1 |
| Query Match: | 96.92% | Indels: | 92 |
| DB: | 6 | Gaps: | 1 |

US-10-077-111-13 (1-384) x AAD45070 (1-1818)

| | | |
|----|-----|---|
| Qy | 1 | MetValIleuLeuIleIsthrIleuAlaAphIleG1YAspAryValAancCyAAlaApe 20 |
| Db | 148 | ATGTGAAGTGTATTCACATTAAGTGTATGAGATGAGATGATCAATGCTGTGCTTC 207 |
| Qy | 21 | SerPheSerIleuA1aThrCySerIleuApyThrlleArlgIeuYrSerIleuAry 40 |
| Db | 208 | TCCTTTCCCTCTTGCTACTGCTCTCTGACAAACAAATGCTGACTGTTACG 267 |
| Qy | 41 | AspPheThrlGluLeuProH1sSerProLeuYrPheH1sthrYrAlaValH1sCyCy 60 |
| Db | 268 | GACTTACTGACATGCGACATCTCTCATTTGAATTCATACCTATCTGCACTGCTG 327 |
| Qy | 61 | CyAPheSerProSerGlyH1sIleuA1aSerCySerThrAspGlyThrThrValIeu 80 |
| Db | 328 | TGTTTCTCCCTTCAGGACATATTTTGGCATGCTGTCAACAGATGATCACTGCT 387 |
| Qy | 81 | TrpAsnThrlGluAangIyIneIleuA1aValMetGluIneProSerGlySerProval 100 |
| Db | 388 | TGGAATACTGAATAATGACAGATGCTGGCAGTGAACAGCTGAGTGGCAGCCTG 447 |
| Qy | 101 | ArgValCyAGIInPheSerProAAsPserThrCyLeuAlaSerGlyValAlaAAsPglYTr 120 |
| Db | 448 | AGGGTTTGGCCAGTTTCCCAAGCTCCACGCTGTGTGATACAGGGCAGCTGATGAAC 507 |
| Qy | 121 | ValValIleuThrAsnAlaGInSerTYrLyLeuYrArCyAGIYSerValIyAspGly 140 |
| Db | 508 | GTCGTTTGTGAATGACAGTCAATCAATTAATATAGATGTGTATGTTAAAGATGG 567 |
| Qy | 141 | SerIleuAlaIaCyAlaPheSerProAangIYSerPheValThrlYSerSerCy 160 |
| Db | 568 | TCCTTGGCGGATGATGATTTCTCTTAATGAAAGCTTCTTGTACTGCTGCTCATGT 627 |
| Qy | 161 | G1YAspIeuThrValItrPAspApyIyMetArCyCyLeuH1sSerGlyValAlaH1sAsP 180 |
| Db | 628 | GGTGATTTAAACAGTGTGGATGATTAATAAGAGTGTCTGATGTGTAATAAGACATGAT 687 |
| Qy | 181 | LeuG1YIleThrCyCyAspPheSerSerGInProValSerAspGlyGluGInG1Yleu 200 |
| Db | 688 | CTTGAAATTAACGTGCTGCGATTTTCTTCAACAGCAGTTTGTATGAGAAACAAGGCTT 747 |

```
OY 201 GlnPhePheAArgLeuAlaSerCysGlyGlnAAspCysGlnValLysIleTyrPheValSer 220
Db 748 CAGTTTTCGACATGGCATCATGTGTCAGATTGCCAAGTCAAAATTGGATTGTTCT 807
OY 221 PheThrIleIleLeu----- 225
Db 808 TTAAACCATATCTTAGGTTTGAATTAAATATATAAGTACAGAGTGGGCACTGTGCT 867
OY 225 ----- 225
Db 868 CTTGTTTGGCTTGTCTTTTCCCATGATGGGCAATGCTAGTCTAGGGTCACTGGAT 927
OY 225 ----- 225
Db 928 AAGTCTGTCAATATATGATGACTATATGAGATAATATCTTCAACATTTGACTCAGCAC 987
OY 225 ----- 225
Db 988 ACCAGTATGTCAACAATGTGCTTTTGGCACTTAATACCTTTTACTTGTACTGTGTCA 1047
OY 226 -----AlaArgArg 228
Db 1048 ATGACAAACAAACAGTGAACATCTGGCAATTGACCTGGAAACCTTTGCCAAGCAAGCGC 1107
OY 229 ThrGlnIleGlnLeuLysGlnPheThrGlnAAspTyrSerGlnGlnValIleSerThrTyr 248
Db 1108 ACAGAAATCATGCTGGAAGCAATTTACCAAGATTGCTCAGAGGAGATGTCTCAACATGG 1167
OY 249 LeuCysAlaGlnAAspLeuLysAAspLeuValGlyIlePheLysMetAAsnIleAspGly 268
Db 1168 CTTGTGTCAACAAGATTAAAGATCTTGTGTGTTTTCATGAAATGAAATGATGATGGA 1227
OY 269 LysGlnLeuLeuAAsnLeuThrLysGlnSerLeuAlaAspAAspLeuLysIleGlnSerLeu 288
Db 1228 AAAAAGACTGTTGAATCTTCAAAAGAAAGTGTGCTATGATGATTTGAAATTTGAAATCTCTA 1287
OY 289 GlyLeuAArgSerLysValLeuArgLysIleGlnGlnLeuAArgThrLysValLysSerLeu 308
Db 1288 GGACTGCGTATTAAGTCTGAGAAATTTGAAAGCTCAGAGACCAAGTTAAATCCCTT 1347
OY 309 SerSerGlyIleProAAspGlnPheIleCysProIleThrAArgLysLeuMetLysAAspPro 328
Db 1348 TCTTCAGGAATTCCTGATGAAATTTATATGTCCAATAGAGAACTTATGAAAGATCCG 1407
OY 329 ValIleAlaSerAAspGlyTyrSerTyrGlnLysGlnValMetGlnAAsnTyrIleSerLys 348
Db 1408 GTCATCCCATCAGATGGCTATTTCAATGAAAGAAAGAAAGAAATTTGATCAGCAAA 1467
OY 349 LysLysAArgThrSerProMetThrAAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db 1468 AAGAAAGCTACAAAGTCCCATACAAATCTTCTTCTTCAAGCGATCTTACACCAAT 1527
OY 369 ArgThrLeuLysMetAlaIleAAsnAArgTyrLeuGlnThrIleGlnLys 384
Db 1528 AGGACTCTGAAATATGGCCATCATATGATGCTGGAGACACCAAAAG 1575

RESULT 4
ABV29028
ID ABV29028 standard; cDNA; 1996 BP.
XX
XX ABV29028;
AC
XX
XX 16-SEP-2002 (first entry)
DT
XX
XX Human prostate expression marker cDNA 29019.
DE
XX
XX Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; 88.
XX
XX Homo sapiens.
OS
XX
XX WO200160860-A2.
XX
```

```
PD 23-AUG-2001.
XX
XX
PF 20-FEB-2001; 2001MO-US005171.
XX
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
PI Schlegel R, Endege WO, Monahan JE;
PI
XX
DR WPI; 2001-662795/76.
XX
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX
PS Claim 1; Page 6145-6146; 11750pp; English.
XX
XX
CC The invention relates to an isolated nucleic acid molecule (1) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (1) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX
SQ Sequence 1996 BP; 615 A; 378 C; 447 G; 543 T; 0 U; 13 Other;
XX
XX
Alignment Scores:
Pred. No.: 4.09e-185 Length: 1996
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 80.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: Gaps: 1

US-10-077-111-13 (1-384) x ABV29028 (1-1996)
OY 1 MetValLysLeuIleIleThrLeuAlaAspHisGlyAAspAAspValAAspCysAlaPhe 20
Db 16 ATGATGAACCTGATTCACACATTAAGCTGATCAAGTACATGCTGTCCTTC 75
OY 21 SerPheSerLeuLeuAlaThrCysSerLeuAAspLysThrIleAArgLeuTyrSerLeuArg 40
Db 76 TCCTTTTCCTCTTGGTACTTGTCTGTCGACAAAACAATTTGGCTGATCTGTTCGT 135
OY 41 AspPheThrGlnLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysGys 60
Db 136 GACTTTACTAAGCTGACACATTTCTCCATTAAGTTTCAATCAATGCTGTCCTGCTGC 195
OY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAAspGlyTyrThrValLeu 80
Db 196 TGTTCCTCCCTTCAGACATATTTTGGCATCTGTTTCAACAATGTATCCACTGTCTTA 255
OY 81 TrpAAsnThrGlnAAsnGlyGlnMetLeuAlaValMetGlnLysProSerGlySerProVal 100
Db 256 TGGAAATATCAAAATGACAGATGCTGGCAGTATGGAACAGCTAGTGGCAGCCCTGTG 315
OY 101 ArgValCysGlnPheSerProAAspSerThrCysLeuAlaSerGlyAlaIleAAspGlyThr 120
Db 316 AGGATTGGCCAGTTTCCCAAGCTCCACCTGTTTGGCATCAGAGGGGCGATGGAAC 375
OY 121 ValValLeuThrAAsnAlaGlnSerTyrLysLeuTyrAArgCysGlySerValLysAAspGly 140
```



```
Db      ||| 376 GTGGTTTGTGGATGACAGTCATCAAAATTATATGATGTGTATGTTAAAGTGC 435
Qy      ||| 141 SerLeuAlaIaIaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCy 160
Db      ||| 436 TCCTTGGCGGCACTGATTTTCTCTTAATGAAAGCTTCTTGTCACTGCTCTCATGT 495
Qy      ||| 161 GlyAspLeuThrValTTPaAspAspLysMetArgCysLeuHisSerGlyValAsp 180
Db      ||| 496 GGTGATTTAAACAGTGTGGATGATTAATAGAGGTGTCTGATGAAAAAGCACAATGAT 555
Qy      ||| 181 LeuGlyTLeThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
Db      ||| 556 CTGGAAATTACCTGCTGCGATTTTCTTTCACAGCAAGTTCTGATGGAAGAACAAGCTT 615
Qy      ||| 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTPriLeValSer 220
Db      ||| 616 CAGTTTTTTGCACTGGCATCATGTGTGATGAGATTGCCAAGTCAAAAATTGGATTGTTCT 675
Qy      ||| 221 PheThrHisIleLeu----- 225
Db      ||| 676 TTTAACCATATCTTAGGTTTGAATTAAATATATAAGTACACTGAGTGGGCACTGTGCT 725
Qy      ||| 225 ----- 225
Db      ||| 736 CCTGTCTGCGCTTGCTGCTTTTCCCATGATGGCGAGATGCTAGTCAAGGTCAGTGAT 795
Qy      ||| 225 ----- 225
Db      ||| 796 AAGTCTGTCTAGTATATGATATCAATATCTGAGAATATCTTACACATTTGACTCAGAC 855
Qy      ||| 225 ----- 225
Db      ||| 856 ACCAGATATGCACAACTTGCTTTTGCACCTAATACCTTTTACTTGCTAGTGTCA 915
Qy      ||| 226 ----- 228
Db      ||| 916 ATGACAAAAACAGTGAACATCTGCAATTGACCTGMAACACTTGTCCAAAGCAAGCGC 975
Qy      ||| 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValValSerThrTrp 248
Db      ||| 976 ACGAACAATCAGCTGAAGCAATTTACGAAAGATTGGCAAGAGAGATGTCTCAACATGG 1035
Qy      ||| 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
Db      ||| 1036 CTTTGTGCAACAAGTTTTAAAAAGATCTTGTGATTTTCAAGATGATTAACATTGATGGA 1095
Qy      ||| 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
Db      ||| 1096 AAAGAAGCTGTGAATCTTACAAAAAGAAAGTGGCTGATGATTTGAAAAATTGAATCTCTA 1155
Qy      ||| 289 GlyLeuAspSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeu 308
Db      ||| 1156 GGAAGTGGTGTAAAGTGAAGGAAAAATTGAAGAGTCAAGAACAAAGTTAAATCCCTT 1215
Qy      ||| 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
Db      ||| 1216 TCTTCAGAGATTCTGTGATTAATTAATGTCCATTAATGAGAACTTAAGAAAGATCCG 1275
Qy      ||| 329 ValIleAspSerAspGlyLysSerTyrgLysGluLysIleMetGluAsnTrpIleSerLys 348
Db      ||| 1276 GTCATCGCATCAGATGCTATTATCATGAAAGAACAAATGGAATAATGGATCAGCAAA 1335
Qy      ||| 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValIleThrProAsn 368
Db      ||| 1336 AAGAAAGTACAAATCCCATGACAAATCTTGTCTTCCACAGGATCACTTAACCAAAAT 1395
Qy      ||| 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGlnThrHisGlnLys 384
Db      ||| 1396 AGGACTGTGAATAATGCCATCAATAGATGCTGGAGACACACCAAAAG 1443
```

```
Id      ABV23190 standard; cDNA; 1996 BP.
Xx      AC      ABV23190;
Xx      DT      16-SEP-2002 (first entry)
Xx      DE      Human prostate expression marker CDNA 23181.
Xx      KW      Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
Xx      KW      pharmacogenomic marker; gene; ss.
Xx      OS      Homo sapiens.
Xx      PN      WO200160860-A2.
Xx      PD      23-AUG-2001.
Xx      PF      20-FEB-2001; 2001MO-US005171.
Xx      PR      17-FEB-2000; 2000US-018319P.
Xx      PR      16-MAR-2000; 2000US-0189862P.
Xx      PR      25-MAY-2000; 2000US-0207454P.
Xx      PR      09-JUN-2000; 2000US-0211314P.
Xx      PR      18-JUL-2000; 2000US-0219807P.
Xx      PR      13-DEC-2000; 2000US-0235281P.
Xx      PA      (MIL-) MILLENNIUM PREDICTIVE MEDICINE INC.
Xx      PI      Schlegel R, Endege WO, Monahan JB;
Xx      DR      WPI, 2001-662795/76.
Xx      PT      Novel isolated nucleic acid molecule associated with cancerous state of
Xx      PT      prostate cells and correlating with presence of prostate cancer, useful
Xx      PT      for detecting presence of prostate cancer, stage of prostate cancer.
Xx      PS      Claim 1; Page 4166-4167; 11750pp; English.
Xx      CC      The invention relates to an isolated nucleic acid molecule (1) comprising
Xx      CC      a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
Xx      CC      specification or its complement. (1) is useful for: (a) assessing whether
Xx      CC      a patient is afflicted with prostate cancer; (b) monitoring the efficacy
Xx      CC      of a test compound to inhibit prostate cancer in a patient; (c) assessing
Xx      CC      the efficacy of a therapy for inhibiting prostate cancer in a patient;
Xx      CC      (d) selecting a composition for inhibiting prostate cancer in a patient;
Xx      CC      (e) assessing the prostate cell carcinogenic potential of a compound; (g)
Xx      CC      determining whether prostate cancer has metastasized in a patient; (h)
Xx      CC      assessing the aggressiveness or indolence of prostate cancer in a patient
Xx      CC      ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
Xx      SQ      Sequence 1996 BP; 615 A; 378 C; 447 G; 543 T; 0 U; 13 Other;

Alignment Scores:
Pred. No.: 4, 09e-185
Score: 1996
Percent Similarity: 1984.00
Best Local Similarity: 80.468
Query Match: 80.468
DB: 96.928
Indels: 92
Gaps: 1

US-10-077-111-13 (1-384) x ABV23190 (1-1996)

Qy      1 MeValLysLeuLeuLeuIleThrLeuAlaAspHisGlyAspAspValAsnCyCysAlaPhe 20
Db      16 ATGGTAAACTGATTACATTAGCTGATCATGAGAGATGTCACATGCTGTGCTTC 75
Qy      21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrsSerLeuArg 40
Db      76 TCCTTTTCCCTCTTGAGCTACTTCTCTTGGACAAACAATTCGCTGTACTGTTACGT 135
Qy      41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyraIaValHisCyCys 60
```


Db 136 GACTTACTGAACTCCACATTCCTCATGTGAAGTTTCATACCTATGCTGCTCAGCTGTC 195
 Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
 Db 196 TGTTCCTCCCTTCCAGACATATTTTGGCATCGTGTTCACACAGATGGTACACAGCTCTTA 255
 Qy 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
 Db 256 TGGAAATCTGAAAATGACAGATCTGCGACAGTGAAGAACGCTAGTGGAGCCCTGTG 315
 Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAspGlyThr 120
 Db 316 AGGGATTGCGCAGTTTCCCGACACTCCAGCTGTTCATCAGGGGACGCTGAAGGAACT 375
 Qy 121 ValValLeuTrpAsnAlaGlnSerTrpLeuValArgCysGlySerValIleAspGly 140
 Db 376 GTGGTTTGTGGATGACAGCTCATACAAATTATATGATGGTAGTAAAGATGGC 435
 Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPheValThrGlySerSerCys 160
 Db 436 TCCTTGGCGGCGATGTGATTTTCTCTAATGAAAGCTTCTTGTCACTGGCTCTCATGT 495
 Qy 161 GlyAspLeuThrValITTPAspAspLysMetArgCysLeuHisSerGlyIleValHisAsp 180
 Db 496 GGTGATTTTAAACAGTGTGGATGATAAATAGGTGTCTGCTAGTGAAGAAAGCACATGAT 555
 Qy 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
 Db 556 CTGGAAATACCTCTCGATGATTTTCTTCACAGCAGAGTTCTGTATGAGAAACAAGGCTT 615
 Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIleTrpIleValSer 220
 Db 616 CAGTTTTTGCACCTGGCATCATGTGTCCAGATTGCCAAGTCAAAATTTGGATTGTTCT 675
 Qy 221 PheThrHisIleLeu----- 225
 Db 676 TTATCCATATCTTAGGTTTGAATTAATAATATAAGTACACTGAGTGGGCACTGTGCT 735
 Qy 225 ----- 225
 Db 736 CCTGTCTGCTGTGCTGCTTTTCCCATGATGGGAGATGCTAGTCTGAGGGTCACTGAT 795
 Qy 225 ----- 225
 Db 796 AAGTGTGATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 855
 Qy 225 ----- 225
 Db 856 ACCAGTATGTCACAACTGTGCTTTTGGACCTAATACCTTTTACTTGTACTGATCA 915
 Qy 226 -----AlaArgArg 228
 Db 916 ATGACAAACAGTGAACATCTGGCAATTGACCTGAAACACTTTGCCAAGCAAGGCGC 975
 Qy 229 ThrGlnIleGlnLeuValPheThrGluAspTrpSerGluGluValIleValSerThrTrp 248
 Db 976 ACAGAAATCAGCTGGAAGCAATTTACCGAAGATTGGTCAAGGAGATGTCTCAACATGG 1035
 Qy 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnIleAspGly 268
 Db 1036 CTGTGTGACAAAGATTTAAAGATCTGTGTGATTTTCAAGATGAATTAACATTTGATGGA 1095
 Qy 269 LysGluLeuLeuAsnLeuThrLysGluSerIleAlaAspLeuValIleGluSerLeu 288
 Db 1096 AAAAAGCTGTGATCTTCAAAAGAAAGTCTGCTGATGATTTTGAATAATGATCTCTTA 1155
 Qy 289 GlyLeuArgSerLysValIleuArgLysIleGluGluLeuArgThrLysValIleSerLeu 308
 Db 1156 GGACTGCGTAAAGAGCTGAGGAAATGAGAAAGCTCAGAGCAAGGTTAAATCCCTT 1215
 Qy 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
 Db 1216 TCTTCAGGAATTCCTGATGAATTTATATGTCATTAATCTAGAGAACTTATGAAAGATCCG 1275

Qy 329 ValIleAlaSerAspGlyTrpSerTrpGluLysGluAlaMetGluAsnTrpIleSerLys 348
 Db 1276 GTCTATGCTATCAATATGCTATTTCAATATGAAAAGAAAGCAATGAAATTTGATCAGCAA 1335
 Qy 349 LysLysArgTrpSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
 Db 1336 AAGAAAGCTACAGTCCCATGACAAATCTTGTCTTCTTCCCTTCAAGCGGTACTTACACCAAAT 1395
 Qy 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys 384
 Db 1396 AGGACTCTGAAAATGGCCATCAATAGATGGCTGGAGACACCAAAAG 1443
 RESULT 6
 ID ABZ24707 standard; cDNA; 1773 BP.
 AC ABZ24707;
 DT 07-APR-2003 (first entry)
 XX
 DE Human cell growth, differentiation and death protein CGDD-19 cDNA.
 XX
 KW 'CGDD-19'; cell growth; cell differentiation; cell death; human;
 KW cytoskeletal; antiarteriosclerotic; hepatotropic; antiinflammatory;
 KW antiporiatic; antianemic; ophthalmological; auditory; anticonvulsant;
 KW cerebroprotective; nocitropic; neuroprotective; antiparkinsonian;
 KW neuroleptic; tranquilizer; immunosuppressive; anti-HIV; antiallergic;
 KW antidiabetic; antihypertoid; antidiabetic; dermatological; nephrotropic;
 KW antineumatic; antiarthritic; antilicer; vulnary; virucide;
 KW antibacterial; fungicide; antiparasitic; protozoicide; antihelminthic;
 KW antifertility; gynaecological; gene therapy; microarray; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 60..1490
 FT /*tag= a
 FT /product= "Human CGDD-19"
 XX
 WO200297032-A2.
 XX
 PD 05-DEC-2002.
 XX
 PF 05-APR-2002; 2002MO-US011152.
 XX
 PR 06-APR-2001; 2001US-0282110P.
 PR 11-APR-2001; 2001US-0283294P.
 PR 26-APR-2001; 2001US-0286820P.
 PR 27-APR-2001; 2001US-0287228P.
 PR 16-MAY-2001; 2001US-0291622P.
 PR 18-MAY-2001; 2001US-0291646P.
 PR 25-MAY-2001; 2001US-0293727P.
 PR 01-JUN-2001; 2001US-0295263P.
 PR 01-JUN-2001; 2001US-0295340P.
 PR 15-JUN-2002; 2002US-0349705P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Aizimzai Y, Au-Young JK, Batra S, Baughn MR, Becha SD;
 PI Boroweky ML, Burford N, Ding L, Elliott VS, Emerling BW, Gandhi AR;
 PI Gietzen KJ, Griffin JA, Hafalala AJA, Honchell CD, Lal PG, Lee SY;
 PI Lu DM, Arvizu CS, Ramkumar J, Reddy R, Sanjanwala MM, Tang YT;
 PI Walla NK, Wang YS, Warren BA, Xu Y, Yang J, Yao MG, Yue H;
 PI Zebajadian Y;
 XX
 DR WPI: 2003-140453/13.
 DR P-PSDB; ABP58348.
 XX
 PT Novel human proteins associated with cell growth, differentiation and
 PT death, useful for treating, diagnosing or preventing cancer,
 PT developmental, neurological, reproductive or autoimmune/inflammatory
 PT disorders.

XX Claim 5, Page 235-236; 238bp; English.

CC The present sequence is that of Incyte clone 7483131CB1 encoding human
CC CGDD-19, a novel protein associated with cell growth, differentiation and
CC death. A representative cDNA library for the polynucleotide is KIDNONT19
CC derived from kidney tissue. Structural features establish the encoded
CC protein as being associated with cell growth, differentiation and death,
CC and as showing sequence and structural similarity to human apoptotic
CC protease activating factor 1. The invention is based on novel human CGDD-
CC 1 to -21 proteins (see ABP58330-50), the polynucleotides encoding them
CC (see AB24689-709), and to the use of these for the diagnosis, treatment
CC or prevention of cell proliferative disorders including cancer,
CC developmental disorders, neurological disorders, autoimmune disorders,
CC reproductive disorders, and disorders of the placenta, and in the
CC assessment of the effects of exogenous compounds on the activity and
CC expression of proteins and nucleic acids associated with cell growth,
CC differentiation and death. CGDD polynucleotides are also used in a
CC claimed microarray and in a claimed method of generating an expression
CC profile of a sample

XX Sequence 1773 BP; 535 A; 334 C; 371 G; 533 T; 0 U; 0 Other;

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 1.35e-184 | Length: | 1773 |
| Score: | 1978.00 | Matches: | 382 |
| Percent Similarity: | 80.25% | Conservative: | 0 |
| Best Local Similarity: | 80.25% | Mismatches: | 2 |
| Query Match: | 96.63% | Indels: | 92 |
| DB: | 8 | Gaps: | 1 |

US-10-077-111-13 (1-384) x AB224707 (1-1773)

QY 1 MetVallyLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe 20
DB 60 ATGTGAAACTGATTCACACATTAGCTGATCATGTCAGATGTCACATCTGCTGCTTC 119
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTYrSerLeuArg 40
DB 120 TCTTTTCCCTCTTGCTACTTCTCTTGGACAAACAATTGCGCTGACTGTTACGT 179
QY 41 AspPheThrGluLeuProHisSerProLeuLeuPheHisThrTYrAlaValHisGlyCys 60
DB 180 GACTTACTGAACTGCACATCTCCATTGAAAGTTTCATACATGATGTCACCTGTC 229
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
DB 240 TGTTCCTCCCTCCAGACATATTTGGCATCGTGTTCACACATGTCACCTGCTCA 259
QY 81 TyrAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
DB 300 TGGAAATACGAAATGACAGATGTCGAGATGGAACAGCCTAGTGGCACCCCTGTC 359
QY 101 ArgValCysGlnInPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr 120
DB 360 AGGATTTGGCCAGTTTCCCAAGCTCCACGCTGTTGGCAACAGGGGACGCTGATGAACT 419
QY 121 ValValLeuThrAsnAlaGlnSerTYrLysLeuTYrArgCysGlySerValLysAspGly 140
DB 420 GTGGTTTGTGAAATGACAGTCAATACAAATTATATGATGTGTGTAAAGATGGC 479
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
DB 480 TCCTTGGCGGCGATGTGATTTTCTCTTAATGGAAGCTTCTTGTCACTGGCTCTCATGT 529
QY 161 GlyAspLeuThrValTTPAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
DB 540 GGGATTTTAAACAGTGTGGATGATTAATAGAGGTCTGATGTGAAAAAGCACTGAT 599
QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
DB 600 CTTGGAAATTACCTGCTGCGATTTTCTTTCACAGCAGTTTCTGATGGAACAAGGCTT 659

QY 201 GlnPheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTTP1IleValSer 220
DB 660 CAGTTTTTTCAGCTGGCATCATGTGGTCAGATTCAGATTCAGAAATTTGGATTTGTTCT 719
QY 221 PheThrHisIleLeu----- 225
DB 720 TTTTACCAATATCTTAGCTTTTGAATTAATAATTAAGTACATGAGTGGGACCTGTCT 779
QY 225 ----- 225
DB 780 CCGTTCTGGCTTGTGCTTTTCCATGATGGGACAGATGCTAGTCTCAGGTCAGTGAAT 839
QY 225 ----- 225
DB 840 AAGTCTGATGATATATGATATCTAATACTGAAATATCTTACACATTGACTCAGCAC 899
QY 225 ----- 225
DB 900 ACCAGGTATGTCAAACTTGTGCTTTTGACCTTAATACCTTTTACTTGTCTGCTTCA 959
QY 226 -----AlaArgArg 228
DB 960 ATGGACAAACAGTGAACATCTGGCAATTGACCTGGAACAATTGGCCACAGGAGAC 1019
QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTTPSerGluGluValLysThrTTP 248
DB 1020 ACAGAACATCAGCTGAAAGCAATTTACCGAAGATGCTCAGAGAGAGATCTTCAACATGG 1079
QY 249 LeuCysAlaGlnAspLeuAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
DB 1080 CTTTGTGCAAGATTTTAAAGATCTTGTGTTATTTTCAAGATGAATACATTGATGAA 1139
QY 269 LysGluLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGlnSerLeu 288
DB 1140 AAAGAACCTTTGATCTTCAABAAGAACTCGGCTGATGATTTGAAATTTGAATCTCTCA 1199
QY 289 GlyLeuAspSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeu 308
DB 1200 GGACTGCTGATGAAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGCTTAAATCCCTT 1259
QY 309 SerSerGlyIleProAspGluPheLysCysProIleThrArgGluLeuMetLysAspPro 328
DB 1260 TCTTCAGGAATTCGATGAATTTATATGTCCAATACATGAACTTAGAACTTAGAAAGTCCG 1319
QY 329 ValIleAlaSerAspGlyTYrSerTYrGluLysGluAlaMetGluAsnTTP1IleSerLys 348
DB 1320 GTCATGCAATCAGATGCTATTCATATGAAAGAAAGCAATGGAATTTGATCAGCACA 1379
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
DB 1380 AAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTCCCTTCAGCGTACTTACACCAAT 1439
QY 369 ArgThrLeuLysMetAlaIleAsnArgTTPLeuGluThrHisGlnLys 384
DB 1440 AGGACTCTAATAATGGCCATCAATAGATGGCTGAGACACACCAAAAG 1487
RESULT 7
ID ADB63535 standard; cDNA; 1811 BP.
AC ADB63535;
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA encoding clone TEST120264530.
XX
KW Human; ss; gene; pharmaceutical; diagnostic; gene therapy;
KW tissue regeneration; cell regeneration; membrane protein;
KW signal transduction-related protein; transcription-related protein;
KW osteoporosis; neurological disease; cancer; tumour.
XX
XX Homo sapiens.
XX

| Key | Location/Qualifiers |
|--|---|
| FT CDS | 150..1580 |
| FT | /*tag= a |
| FT | /product= "Clone TEST120264530 protein" |
| XX | |
| PN | EP1308459-A2. |
| XX | |
| PD | 07-MAY-2003. |
| XX | |
| XX | 28-MAR-2002; 2002EP-00007401. |
| XX | |
| PR | 05-NOV-2001; 2001JP-00379298. |
| PR | 25-JAN-2002; 2002US-00350976. |
| XX | |
| PA | (HELI-) HELIX RES INST. |
| PA | (REAS-) RES ASSOC BIOTECHNOLOGY. |
| PI | Isegaki T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Iehi S; |
| PI | Yamamoto J, Isono Y, Hio Y, Otsuka K, Nagai K, Irie R, Tamechika I; |
| PI | Seki N, Yoshikawa T, Otsuka M, Nagahari K, Masuno Y; |
| XX | |
| XX | WPI; 2003-450961/43. |
| PS | P-PSDB; ADB65505. |
| PT | |
| PT | New polynucleotides and polypeptides, useful for developing a diagnostic |
| PT | marker or medicines for regulation of their expression and activity, or |
| PT | as targets of gene therapy. |
| XX | |
| PS | Claim 1; Page; 222pp; English. |
| XX | |
| CC | The invention discloses a polynucleotide comprising a sequence selected |
| CC | from 1970 fully defined nucleotide sequences which encode novel |
| CC | polypeptides. Also claimed is a polypeptide encoded by the polynucleotide |
| CC | or its partial peptide, an antibody binding to the polypeptide or peptide |
| CC | of the polynucleotide, immunologically assaying the polypeptide or |
| CC | peptide of the polynucleotide by contacting the polypeptide or peptide |
| CC | with the antibody of the encoded protein, and observing the binding |
| CC | between the two, a transformant carrying the polynucleotide in an |
| CC | expressible manner and an antisense polynucleotide. The oligonucleotide |
| CC | is useful as a primer for synthesizing the polynucleotide, or as a probe |
| CC | for detecting the polynucleotide. The polynucleotides and encoded |
| CC | proteins are useful as pharmaceutical agents and many disease-related |
| CC | genes may be included in them, for developing a diagnostic marker or |
| CC | medicines for regulation of their expression and activity, or as targets |
| CC | of gene therapy. The genes are involved in tissue and/or cell |
| CC | regeneration. Membrane proteins, signal transduction-related proteins, |
| CC | transcription-related proteins, disease-related proteins and genes |
| CC | encoding them can be used as indicators for diseases (e.g. osteoporosis, |
| CC | neurological diseases, cancer, tumor). The cDNA may be used to regulate |
| CC | the activity or expression of the encoded protein to treat diseases. The |
| CC | sequence presented is a cDNA of the invention. Note: Some of the sequence |
| CC | data for this patent is not represented in the printed specification, but |
| CC | is based on sequence information supplied by the European Patent Office. |
| XX | |
| SQ | Sequence 1811 BP; 520 A; 352 C; 399 G; 540 T; 0 U; 0 Other; |
| XX | |
| Alignment Scores: | |
| Pred. No.: | 1,39E-184 |
| Score: | 1978.00 |
| Percent Similarity: | 80.25% |
| Best Local Similarity: | 80.25% |
| Query Match: | 96.63% |
| DB: | 10 |
| | Gaps: 1 |
| US-10-077-111-13 (1-384) X ADB63535 (1-1811) | |
| 0Y | 1 MetValIyleLeuIleIleIsthrIleuAlaAPHIsgIyAspAspValaIsmCySalaIaphe 20 |
| Db | 150 ATGGTGAACCTGATTCAACATTAGCTGATTCATGTCAGTACGATGCAACTGCTGTCCTTC 205 |
| 0Y | 21 SerPheSerIleuLeuIleIamThrCySeriIaAspIySthIleArgLeuTySeriLeuArg 40 |
| Db | 210 TCGCTTTCCTCTTGCTGCTACCTGCTCTTGACCAAAACAAATTCGGCTGACTCGTTACGT 265 |

| | | | |
|----|------|--|------|
| QY | 41 | AspPheThrGluLeuProHisSerProLeuIysPheHisThrTyrAlaValHisCysGys | 60 |
| DB | 270 | GACTTACTGAACCTGCCACATCTTCCATGTGAAGTTTCAATACCTTACTGCTTCACATGCTGC | 329 |
| QY | 61 | CysPheSerProSerGlyHisIleLeuAlaSerCysSerThraSpGlyThrThrValIleu | 80 |
| DB | 330 | TGTTTTCCTCCCTTCAGGACATATTTTGGCATCGTGTTCAACAGATGGTACCACTGTCTTA | 389 |
| QY | 81 | TrpAenThrGluAenGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal | 100 |
| DB | 390 | TGGAATACCTGAAATATGACAGATCTGGCAGGTATGGAACAGCCTTAGGCGACGCTGTG | 449 |
| QY | 101 | ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaIaAspGlyThr | 120 |
| DB | 450 | AGGGTTTGGCCAGTTTCCCGACACTCCACGGTTTGGCATCAGGAGGACGCTGATGAATCT | 509 |
| QY | 121 | ValValLeuTrpAenAlaGlnSerTyrIleLeuTyrArgCysGlySerValIysAspGly | 140 |
| DB | 510 | GTCGTTTGTGGAATGACAGCTCATCAAAATTAATAGATGTGTAGTGTAAAGATGCG | 569 |
| QY | 141 | SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValIlnGlySerSerCys | 160 |
| DB | 570 | TCCTTTGGGGGATGGGATTTTCTCTTAATGAAGCTTCTTTGTACATGGCTCTCATGT | 629 |
| QY | 161 | GlyAspLeuThrValIleTrpAspAspLysMecArgCysLeuHisSerGlyLysAlaHisAsp | 180 |
| DB | 630 | GGTACATTTAACAGTGTGGGATGATAAATAGGGTGTCTGCATATGTGAAAAAGCACATGAT | 689 |
| QY | 181 | LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu | 200 |
| DB | 690 | CTTGAAATATACCTCTCGGATTTTCTTCACAGCGCAAGTTTGTAGTAGGAACAGGCTT | 749 |
| QY | 201 | GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIysIleTrpIleValSer | 220 |
| DB | 750 | CAGTTTTTTTGCACCTGGCATCATGTGTGAGGATTTGCCAAGTCAAAATTTGGATTGTTCT | 809 |
| QY | 221 | PheThrHisIleLeu----- | 225 |
| DB | 810 | TTTACCCCATATCTTAGGTTTGAATTTAAATATATAAAGTACACTGAGTGGGCACGTGCT | 869 |
| QY | 225 | ----- | 225 |
| DB | 870 | CCTGTTCTGGCTTGCTTTTCCATGATGGGAGATGCTAGTCTCAGGGTCAGTGAT | 929 |
| QY | 225 | ----- | 225 |
| DB | 930 | AAAGTCTGTCATATATATGATATCTAATACTGAGAAATATATCTTCACACTTACATCAGAC | 989 |
| QY | 225 | ----- | 225 |
| DB | 990 | ACCAAGTATGTCACAACTTGCTTTGGACCTAATACCCTTTACTGTACTGCTTCA | 1049 |
| QY | 226 | -----AlaArgArg | 228 |
| DB | 1050 | ATGGAACAACACGTGAACATCTGGCAATTTGACCTGGAAACACTTTGGCCAAAGCAAGAGC | 1109 |
| QY | 229 | ThrGlnHisGlnLeuLysGlnPheThrGluAspTrpSerGluGlnValValSerThrTrp | 248 |
| DB | 1110 | ACGAAACATCGCTGAGAACAAATTTTACCGAAATTTGTGTGACGAGAGATGTTCTTCACATGG | 1168 |
| QY | 249 | LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly | 268 |
| DB | 1170 | CTTTGTGACAAAGATTTTAAAGATCTTGTGTATTTTTCACAGATGAATTAACATTGATGGA | 1229 |
| QY | 269 | LysGlnLeuLeuAenLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGlnSerLeu | 288 |
| DB | 1230 | AAAGAACTGTGAATCTTACAAAGAAAGAAAGTCTGCTGATGATTTGAAAAATTTGAATCTCTTA | 1288 |
| QY | 289 | GlyLeuLeuArgSerLysValIleuArgLysIleGlnGlnLeuArgThrLysValIysSerLeu | 308 |
| DB | 1290 | GGACTGCGTATTAAGTCTGAGGAAATTTAAGAGCTCAGAGACCAAGTTTAAATTCCTT | 1349 |

QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuWetIleAspPro 328
DB 1350 TCTTCAGAGATTCCTGATGAAATTAATGTCCAATACTAGAGAACTTATGAAAGATCCG 1409
QY 329 ValIleAlaSerAspGlyIleProAspGluPheIleCysProIleThrArgGluLeuWetIleAspPro 348
DB 1410 GTCATGCAATCAATGCTATTCATATGAAAGAGAACCAATGAAATTTGATCAGCAAA 1469
QY 349 LysLysArgThrSerProMetThrAsnLeuValIleuProSerAlaValIleuThrProAsn 368
DB 1470 AAGAAAGTACAAAGTCCCATGACAAATCTTCTTCTCAGGGGTACTTACACCAAT 1529
QY 369 ArgThrLeuLysMetAlaIleAsnArgTyrLeuGluThrHisGlnLys 384
DB 1530 AGGACTCTGAATAAGGCCATCAATAGATGGCTGGAGACACCAAAAG 1577
RESULT 8
ADB47507
ID ADB47507 standard; cDNA; 1817 BP.
XX
AC ADB47507;
XX
DT 04-DEC-2003 (first entry)
XX
DE Human cDNA upregulated in dendritic cells SEQ ID NO 207.
XX
KM aa; gene; human; dendritic cells; high throughput; cancer;
KM infectious disease; autoimmune disease; allergy;
KM graft versus host disease; vaccine enhancing; gene therapy.
XX
OS Homo sapiens.
XX
PN US2003134283-A1.
XX
PD 17-JUL-2003.
XX
PF 03-OCT-2001; 2001US-00971392.
XX
PR 03-OCT-2000; 2000US-0237652P.
XX
PA (PETE/) PETERSON D P.
PA (PEAR/) PEARSON C I.
PA (COCK/) COCKS B G.
XX
PI Peterson DP, Pearson CI, Cocks BG;
XX
DR WPI; 2003-662509/62.
XX
PT New combination comprises cDNAs that are differentially expressed in
PT dendritic cells useful for preparing a composition for diagnosing or
PT treating cancer, infectious disease, autoimmunity, allergy or graft
PT versus host disease.
XX
PS Claim 1; SEQ ID NO 207; 28bp; English.
XX
CC The invention relates to a combination comprising cDNAs that are
CC differentially expressed in dendritic cells (DC). Also included is a high
CC throughput method for detecting differential expression of one or more
CC cDNAs in a sample containing nucleic acids. The combination is useful for
CC preparing a composition for diagnosing, treating and monitoring the
CC treatment of cancer, infectious disease, autoimmunity, allergy or graft
CC versus host disease, or for enhancing a vaccine. The present sequence
CC represents a human cDNA upregulated in dendritic cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html?docID=20030134283.
XX
SQ Sequence 1817 BP; 519 A; 373 C; 383 G; 542 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1.39e-184 Length: 1817
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0

Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 10 Gaps: 1
US-10-077-111-13 (1-384) x ADB47507 (1-1817)
QY 1 MetValIleLeuLleHisThrLeuAlaAspHisGlyAspAspValAsnCysGlyAlaPhe 20
DB 148 ATGGTGAACCTAATTCACACATTAGCTGATCGTGAAGAGATGCAACCTGCTGCTTC 207
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40
DB 208 TCTTTTCCCTCTTGCTCTCTTGCTCTTGCAAAACAATTCGCTGACTCTTACCT 267
QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrLysAlaValIleCysCys 60
DB 268 GACTTACTGAACCTGACACATTCATGAAAGTTTCAACCTGACCTGCTCAGCTGCTC 327
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValIleu 80
DB 328 TGTTCCTCCCTTCAGACATATTTTGGCATGTCTTCAACAGATGTGCCACTGCTCA 387
QY 81 TrpAsnThrGluAsnGlyIleMetLeuAlaValMetGluGlnProSerGlySerProVal 100
DB 388 TGGAAATACGAAATGACAGATGCTGGCAGATGAGAGAGCTTATGCGACGCTGTG 447
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
DB 448 AGGGTTGGCAAGTTTCCCAAGCTCCAGCTGTTGGATCAGGGGCACTGATGGAAT 507
QY 121 ValIleuThrAsnAlaGlnSerTyrLysLeuTyrArgCysGlySerValIleAspGly 140
DB 508 GTGGTTTGTGAAATCAGATCATCAATTAATATATGATGTGTATTAAGATGCG 567
QY 141 SerLeuAlaAlaCysValaPheSerProAsnGlySerPheValThrGlySerSerCys 160
DB 568 TCTTGGCGCAGATGATTTCTCTTAATGAGCTTCTTGTGCACTGCTCTCATATG 627
QY 161 GlyAspLeuThrValIleTrpAspAspLysMetArgCysLeuHisSerGluLysAlaHisAsp 180
DB 628 GGTGATTTAAACAGTGGGATGATTAATAAGAGTGTCTGCATAGTAATAAGACACTGAT 687
QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
DB 688 CTGGAAATTAACCTGCTGCGATTTTCTTCAAGCCGATTTCTGATGAGAAACAAGCTCT 747
QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIleThrIleValSer 220
DB 748 CAGTTTTCGACTGGCATCATGTGTCAAGATGCCAAATTTGGATGTGTTCT 807
QY 221 PheThrHisIleLeu----- 225
DB 808 TTTTACCATATCTTAGGTTTGAATTAATAATTAAGACACTGAGTGGCACTGTGCT 867
QY 225 ----- 225
DB 868 CTTGTTGCGCTTGCTCTTTTCCATGATGGCAGATGCTAGTCTCAGGGTCAAGTAT 927
QY 225 ----- 225
DB 928 AAGTCTGTCAATAGTATATGATACTAAATGAGAAATATACCTTCTTACTGCTACTGTTCA 987
QY 225 ----- 225
DB 988 ACCAGTATGTCAACAATGTTGCTTTTGACCTTAATCCCTTTTACTGCTACTGTTCA 1047
QY 226 -----AlaArgArg 228
DB 1048 ATGAGCAAAACAGTGAACATCTGGCAATTGACCTGAAACAATTGGCCAGAAAGAGAC 1107
QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTyrSerGlyValIleSerThrTrp 248
DB 1108 ACAGAACATCAGCTGAAGCAATTTACCGAAGATTGCTCAGAGAGAGATGTCTCAACATGG 1167

| | | | |
|----------|---|---|------|
| Qy | 249 | LeuCYsAlaGlnApLeuLeuLYsAspLeuValGlyIlePheLYsMeAsnAsnIleAspGly | 266 |
| Db | 1168 | CTTGTGCAAGATTATAAAGATCTGTGGATTATTTCAGATGAATTAACATTGATGCA | 1227 |
| Qy | 269 | LYsGluLeuLeuAsnLeuThrLYsGluSerLeuAlaAspLeuLYsIleGluSerLeu | 288 |
| Db | 1228 | AAAGACTGTTGGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAATTGAATCTCTA | 1287 |
| Qy | 289 | GlyLeuArgSerLYsValLeuArgLYsIleGluGluLeuValGlyLYsValLYsSerLeu | 308 |
| Db | 1288 | GGACTGCTAGTAAAGTCTGAGGAAAAATTGAAAGCTCAGAGCAACAAAGTTAAATCCCTT | 1347 |
| Qy | 309 | SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLYsAspPro | 328 |
| Db | 1348 | TCTTCAGGAATTCCTCGATGAATTTATAGTCCAAATTAAGTAAGAACTTATGAAAGATCCG | 1407 |
| Qy | 329 | ValIleAlaSerAspGlyTYrSerTYrGluLYsGluAlaMetGluAsnTrpIleSerLYs | 348 |
| Db | 1408 | GTCATCGCATCAGATGGCTATTCAATGATAAAAGAAAGCAATGAAATTTGGATCAGCAA | 1467 |
| Qy | 349 | LYsLYsAspGlyThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn | 368 |
| Db | 1468 | AAGAAACGTACAAAGTCCCATGACAAATCTTGTTCTTCCATCAGCGGTACTTAACCAAAAT | 1527 |
| Qy | 369 | ArgThrLeuLYsMetAlaIleAsnArgTrpLeuGluThrHisGluLYs | 384 |
| Db | 1528 | AGGACTCTGAAATATGGCCATCATATGATGGCTGGAGACACACCAAAAG | 1575 |
| RESULT 9 | | | |
| AI | AI58876 | | |
| ID | AA158876 | standard; cDNA; 1844 BP. | |
| AC | AA158876; | | |
| XX | | | |
| DT | 22-OCT-2001 | (first entry) | |
| XX | | | |
| DE | Human polynucleotide SEQ ID NO 1079. | | |
| XX | | | |
| Kw | Human; noctropic; immunosuppressant; cytosstatic; gene therapy; cancer; | | |
| Kw | peripheral nervous system; neuropathy; central nervous system; CNS; | | |
| Kw | Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; | | |
| Kw | amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; | | |
| Kw | chemokineic; thrombolytic; drug screening; arthritis; inflammation; | | |
| Kw | leukaemia; ss. | | |
| XX | | | |
| OS | Homo sapiens. | | |
| XX | | | |
| PN | WO200153312-A1. | | |
| XX | | | |
| PD | 26-JUL-2001. | | |
| XX | | | |
| PF | 26-DEC-2000; 2000WO-US034263. | | |
| XX | | | |
| PR | 23-DEC-1999; 99US-00471275. | | |
| PR | 21-JAN-2000; 2000US-00488725. | | |
| PR | 25-APR-2000; 2000US-0052317. | | |
| PR | 20-JUN-2000; 2000US-00588042. | | |
| PR | 19-JUL-2000; 2000US-00620312. | | |
| PR | 03-AUG-2000; 2000US-00653450. | | |
| PR | 14-SEP-2000; 2000US-00662191. | | |
| PR | 19-OCT-2000; 2000US-00693036. | | |
| PR | 29-NOV-2000; 2000US-00727344. | | |
| XX | | | |
| PA | (HYSE-) HYSEQ INC. | | |
| XX | | | |
| PI | Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D, | | |
| PI | Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA, | | |
| PI | Zhou P, Goodrich R, Dzmanac RT; | | |
| XX | | | |
| XX | WPI: 2001-44253/47. | | |
| DR | P-PSDB; AAM39720. | | |
| XX | | | |

| | | |
|--|---|---|
| PT | | Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries. |
| XX | | Claim 1; SEQ ID NO 1079; 10078bp; English. |
| PS | | |
| XX | | The invention relates to human nucleic acid (AA157798-AA161369) and the encoded polypeptides (AA38642-AA442213) with nootropic, |
| CC | | immunosuppressant and cyrostatic activity. The polymucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilization of the activities such as: Immune system suppression, CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, CC assays for receptor activity, arthritis and inflammation, leukemias and CC C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification |
| XX | | |
| SQ | Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other; | |
| Alignment Scores: | | |
| Pred. No.: | 1,42e-184 | Length: 1844 |
| Score: | 1978.00 | Matches: 382 |
| Percent Similarity: | 80.25% | Conservative: 0 |
| Best Local Similarity: | 80.25% | Mismatches: 2 |
| Query Match: | 96.63% | Indels: 92 |
| DB: | 4 | Gaps: 1 |
| US-10-077-111-13 (1-384) x AA158876 (1-1844) | | |
| OY | 1 MetValIysLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCySGValAPhe | 20 |
| Dd | 164 ATGTGAACACTGCATTACACATTTACTGTCATCGTGGAGTGCACTGCTGGCCTTC | 223 |
| OY | 21 SerPheSerIleuLeuAlaThrCySSerIleuAspLysThrIleArgLeuTYrSerLeuArg | 40 |
| Dd | 224 TCCITTTCCCTCTGGCTACTGCTGCTCTTGACAAACAATAATCGCTGTACTCGTTACGT | 283 |
| OY | 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTYraIalaValHisCyGS | 60 |
| Dd | 284 GACTTTAATGSAAGTGCACATTTCTCATTTGAAGTTTACTATACCTAATGCTGCACGTGTC | 343 |
| OY | 61 CySPheSerProSerGlyHisIleIleuAlaSerCySSerThrAspGlyThrThValIeu | 80 |
| Dd | 344 TGTTTTCCCCCTTCACAGAATATTTTGGCATCGTGTTCACAGAATGTAACAACGTGCTTA | 403 |
| OY | 81 TrpAsnThrGluAsnGlyGlnMetIleuAlaValMetGluGlnInProSerGlySerProVal | 100 |
| Dd | 404 TGGAAATCTGAANAATGACACAGATGCTGGCAGTAGTGAACAAGCTTAGTGGCAAGCCTGTG | 463 |
| OY | 101 ArgValCySGlnPheSerProAspSerThrCySLeuAlaSerGlyAlaIalAspGlyThr | 120 |
| Dd | 464 AGGGTTTGGCAGTTTTCCCACAGCTCCACAGTGTGGCATCAGGGGCAAGCTGATGAAC | 522 |
| OY | 121 ValValIeuTrpAsnAlaGlnSerTyrlsLeuTYraGySGlySerValIylAspGly | 140 |
| Dd | 524 GTGGTTTTTGGGAATCACAGTCAATCAAAATTAATTAATGTGGATGTGGATTAAGGAATGCC | 583 |
| OY | 141 SerIleuAlaIaCyValAlaPheSerProAsnGlySerPhePheValThnGlySerSerCyS | 160 |
| Dd | 584 TCCTTGGCGGCAATGCATTTTCTCTTAATGGAAAGCTTCTTGTACACGTGGCTCTCAATGT | 643 |
| OY | 161 GlyAspLeuThrValITrpAspAspLysPheLarGySLeuHisASerGlyulTYraIalHisASP | 180 |
| Dd | 644 GGCGATTTAACAGTGTGGGATGATTAATAATGAGGTGTCTGCATAGTGAAGAAAAGACATGAT | 703 |
| OY | 181 LeuGlylYleThrCyAcGAspAspPheSerSerGlnProvalIserrApGlyGluGlnGlyIeu | 200 |
| Dd | 704 CTGGGAATTAACCGCTGCGCAATTTTCTTCACAGCCAGTTTCTGATGGAGAAACAGGCTTT | 763 |

| | | | | |
|-----------|---|------|--|------|
| QY | | 201 | GLNPHEPEAAGLEALASerCYGLVGLNAPCYSGILNVLYLLETMLLVAISer | 220 |
| Db | | 764 | CAGTTTTTTCACCTGGCAATCATGTGGTGAAGATTCGCAAGTCMAAATTTGGATTGTTCT | 823 |
| QY | | 221 | PheThrHisIleLeu----- | 225 |
| Db | | 824 | TTTACCCATATCTTAGGCTTTGAATTAATAATATAAAAGTACATGAGTGGGCACTGTGCT | 883 |
| QY | | 225 | ----- | 225 |
| Db | | 884 | CCTGTTCTGGCTTGCTTTTCCCATGATGGGAGATGTAGTCTACGGGTCAGTGGAT | 943 |
| QY | | 225 | ----- | 225 |
| Db | | 944 | AAGTCTGATAGTATATATATATACTGAGATATATATCTTACACATGTGACTCAGCAC | 1003 |
| QY | | 225 | ----- | 225 |
| Db | | 1004 | ACCAGGATATGCACACACTTGTGCTTTTGCACTTAATACCTTTTATCTTGCTACCTGGTTCA | 1063 |
| QY | | 226 | -----AlaArgArg | 228 |
| Db | | 1064 | ATGGACAAAACAGTACATCTGGCAATTTGACCTCGAAGAACCTTGGCCAGCAAGAGAGC | 1123 |
| QY | | 229 | ThrGLuHISGILNVLYSGILNPhetHrGLuAspTTPSerGIUGluValIserThrTrp | 248 |
| Db | | 1124 | ACAGAACATCTACGCTGAGAACAAATTTACCAAGATTTGGTCTAGAGGAGATGTCTCAACATGG | 1183 |
| QY | | 249 | LeuCYValIAGIAspLeuLYAspLeuValGIlyIlePheLYsMetAsnAsnIleAspGly | 268 |
| Db | | 1184 | CTTTGTGCACAGATTTTAAAGATCTTTGTGTATTTTCAGATGAATTAACATTTGATGGA | 1243 |
| QY | | 269 | LYSGILNVLeuAsnLeuThrLYSGILNVSerLeuAlaAspAspLeuLYSILEGILNVSerLeu | 288 |
| Db | | 1244 | AAAGAACTGTTGAATCTTACAAAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTA | 1303 |
| QY | | 289 | GLYLeuAspSerLYsValLeuArgLYILEGILNVLeuAspGthrLYsValLYsSerLeu | 308 |
| Db | | 1304 | GGACTGCGTAGTAAAGTCTGAGGAAATTAAGAGAGCTCAGGACCAAGGTTAAATCCCTT | 1363 |
| QY | | 309 | SerSerGIlyIleProAspGluPheIleCYsProIleThrArgGluLeuMetLYAspPro | 328 |
| Db | | 1364 | TCTTCAGGAATTCCTGATGAATTTATATATGTCCAATTAACTAAGAACTTATGAAGATCCG | 1423 |
| QY | | 329 | ValIleAlaSerAspGLYTrpSerTYrGIuLYSGILNVAlaMetGluAsnTrpIleSerLYs | 348 |
| Db | | 1424 | GTCATCGCATCAGATGGCTATTCATATATAAAGAAAGAACATGAGAAATTTGATCAGCAA | 1483 |
| QY | | 349 | LYsLYsAspGthrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn | 368 |
| Db | | 1484 | AAGAAACGTACAGTCCCATGCAATCTTGTTCTTCTTCACGGCGGTACTTATCACCAAT | 1543 |
| QY | | 369 | ArgThrLeuLYsMetAlaIleAsnArgTrpLeuGIuThrHISGILNVLYs | 384 |
| Db | | 1544 | AGGACTCTGAAAAATGGCCATCATATGATGAGTCTGAGACACACCAAAAG | 1591 |
| RESULT 10 | | | | |
| ADQ99098 | | | | |
| ID | ADQ99098 standard; cDNA, 1844 BP. | | | |
| XX | ADQ99098; | | | |
| XX | 23-SEP-2004 (first entry) | | | |
| DE | DNA encoding human GPCR-like protein seqid 768. | | | |
| XX | | | | |
| KV | ophthalmological; immunomodulatory; cytosstatic; antiatherosclerotic; | | | |
| KV | antidiabetic; GPCR-like protein; ophthalmic disorder; | | | |
| KV | neurological disorder; immunological disorder; nephritic disorder; | | | |
| KV | hormonal dysfunction; cancer; atherosclerosis; diabetes; | | | |
| XX | molecular weight marker; food supplement; human; ss. | | | |
| XX | | | | |
| DS | Homo sapiens | | | |

| XX | US659662-B1. |
|----|--|
| XX | 27-MAY-2003. |
| XX | 19-JUL-2000; 2000US-00620312. |
| XX | 21-JAN-2000; 2000US-00488725. |
| XX | 25-APR-2000; 2000US-00552317. |
| XX | (HYSE-) HYSEQ INC. |
| XX | Tang YT, Zhou P, Dermanac RT; |
| XX | WPI, 2001-442255/47. |
| XX | New G-protein-coupled receptor-like polypeptides and polynucleotides, |
| XX | useful for treating diseases of ophthalmic, neurological, immunological |
| XX | and nephritic systems and hormonal dysfunction, cancer, atherosclerosis |
| XX | and diabetes. |
| XX | Example 2; SEQ ID NO 768; 92pp; English. |
| XX | The invention describes an isolated polynucleotide (I) comprising a fully |
| XX | defined (SI) of 749, 3188, 2484, 1169, 2936, 1467, 5773, 5714, 4041, |
| XX | 1372, 3996, 3945, 2735, 1788, 585, 1782, 927, 5714 or 2282 nucleotides as |
| XX | CC given in the specification, its translated or protein coding portion, its |
| XX | extracellular portion or its active domain. The GPCR-like polypeptides |
| XX | and polynucleotides are useful for the treatment of diseases of |
| XX | ophthalmic, neurological, immunological and nephritic systems. They may |
| XX | also be used to treat hormonal dysfunction, cancer, atherosclerosis and |
| XX | diabetes. The antibodies are useful for detecting or quantitating the |
| XX | polypeptide in tissue. The polypeptides can also be used as molecular |
| XX | weight markers and as a food supplement. This sequence represents a human |
| XX | polynucleotide of the invention. |
| XX | Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other; |
| XX | Alignment Scores: |
| XX | Pred. No.: 1,426-184 Length: 1844 |
| XX | Score: 1978.00 Matches: 382 |
| XX | Percent Similarity: 80.25% Conservative: 0 |
| XX | Best Local Similarity: 80.25% Mismatches: 2 |
| XX | Query Match: 96.63% Indels: 92 |
| XX | DB: 5 Gaps: 1 |
| XX | US-10-077-111-13 (1-384) x ADQ99098 (1-1844) |
| XX | 1 MetValLysLeuIleHisIleThrLeuAlaAspHisGlyAspAspValaAsnCysGlyAlaPhe 20 |
| XX | 164 ATGTGGAACATGATTCACACATTAGCTGATCATGTGGAGATGTCAACATGCTGTGCTTC 223 |
| XX | 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40 |
| XX | 224 TCCCTTTCCCTCTGGCTACTGCTCTCTGGACAAACAATTCGCTGTACCTGGTACGT 283 |
| XX | 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrValaHisCysCys 60 |
| XX | 284 GACCTTACTGAACATGCGACATTCCTCCATTGAAGTTTCATACCTATGCTGCCATGCTGC 343 |
| XX | 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80 |
| XX | 344 TCTTCTCCCTTCAGAGACATATTTTGGCATGTGTTCAAACGATGGTACCATGTCCTA 403 |
| XX | 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100 |
| XX | 404 TGGATTAATGAAAAGAGACAGATGCTGGACGATGAGAAACAGCCTATGTGACGCTGTG 463 |
| XX | 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValaHisAspGlyThr 120 |
| XX | 464 AGGGTTTGGCAATTTTCCCAACATCCACGTGTGTTGGCATCAGGGGCACTGATGGAAC 523 |
| XX | 121 ValIleLeuTrpAsnAlaGlnSerTyrIleValLeuTyrArgCysGlySerValIleAspGly 140 |


```

Db      524 GTGGTTTGTGGATGACAGCTCATACAAATTATATAGATGTGGTAGTAAAGATGCG 583
Qy      141 SerLeu1a1a1aCySa1aPheSerProsnGlySerPhePheValThrGlySerSerCys 160
Db      584 TCTTGCGGCGATGCGATTTCTCTTAATGGAAGCTTCTTGTCACGCGGCTCCCTCATGT 643
Qy      161 GlyAspLeuThrVal1TrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
Db      644 GGTATTTTAACAGTGTGGATGATTAATAAGTGTCTGCAATGTAAGAAAAAGCAGTAT 703
Qy      181 LeuGly1IleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlyLeu 200
Db      704 CTGGAAATTACCTGCTGCGATTTTCTTCAAGCCAGTTTCTGATGAGAACAAAGTCTT 763
Qy      201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLys1IleTrp1IleValSer 220
Db      764 CAGTTTTTTCAGCTGCGCATCATGTGTGAGATTGCGCAAGTCAAAATTTGGATTGTTTCT 823
Qy      221 PheThrHis1IleLeu----- 225
Db      824 TTTACCCATATCTTAGTCTTTGAATTAATAATATAAAGTACATGAGTGGCAGCTGTCT 883
Qy      225 ----- 225
Db      884 CCTGTTCTGCGTGTGCTTTTCCCATGATGGCGAGATGTAAGTCTCAGGGTCACTGGAT 943
Qy      225 ----- 225
Db      944 AAGTCTGTATGATATGATATGATCTAATCTGAGATATATCTTCAACATGCTCAGCAC 1003
Qy      225 ----- 225
Db      1004 ACCAGATATGACAACTTGCTTTGCTTGGACCTAATACCCTTTACTTGTCTACTGTTCA 1063
Qy      226 ----- 228
Db      1064 ATGACAAAAACAGTGAACATCTGGCAATTGACCTGAAACACTTTGGCCAGCAAGAGGC 1123
Qy      229 ThrGlnHisGlnLeuLysGlnPheThrGlnAspTrpSerGlnGlnVal1ValSerThrTrp 248
Db      1124 ACAGAACTCTGAGCTGAAGCAATTTTACCAAGATTGGTCTCAGAGGAGATGTCTCAACATGG 1183
Qy      249 LeuCysAlaGlnAspLeuLysAspLeuValGly1IlePheLysMetAsnAsn1IleAspGly 268
Db      1184 CTTTGTGCAACAAGATTAAAGATCTGTGTGATTTTCAAGATGAATTAACATGATGGA 1243
Qy      269 LysGlnLeuLeuAsnLeuThrLysGlnSerLeu1aAspAspLeuLys1IleGlnSerLeu 288
Db      1244 AAAGAACTGTTGATCTTACAAAGAAAGTCTGCTGATGATTTGAAATTTGAATCTCTA 1303
Qy      289 GlyLeuAspSerLysValLeuArgLys1IleGlnGlnLeuArgThrLysValLysSerLeu 308
Db      1304 GGACTGCGCTAGTAAAGTGTCTGAGCAAAATTTGAAGAGCTCAGAGCAAAAGTTAAATCCCTT 1363
Qy      309 SerSerGly1IleProAspGlnPhe1IleCysPro1IleThrArgGlnLeuMetLysAspPro 328
Db      1364 TCTTCAGGAATTCCTGATGAATTTATATGTCAAATTAAGTAACTTAAGAACTTATAAAGATCCG 1423
Qy      329 Val1IleAlaSerAspGlyTrpSerTrpGlnLysGln1aMetGlnAsnTrp1IleSerLys 348
Db      1424 GTCAATCGCATCAGATGCGTATTTCAATATGAAGAAAGAAAGCAATTTGATATGACGAAA 1483
Qy      349 LysLysValArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProsn 368
Db      1484 AAGAAAGCTACAAAGTCCCATATCAAAATCTTGTCTCTTCCCTCAGCGGTACTTACACCAAT 1543
Qy      369 ArgThrLeuLysMetAla1IleAsnArgTrpLeuGlnThrHisGlnLys 384
Db      1544 AGGACTCTGAATAATGGCCATCATATGATGCTGGAGACACACCAAAAG 1591

```

```

ID      ADB48858 standard; cDNA; 1844 BP.
XX
AC      ADB48858;
XX
DT      04-DEC-2003 (first entry)
XX
DE      Novel human cDNA seq ID NO 768.
XX
KW      ser; cancer; neurodegenerative disease; human.
XX
OS      Homo sapiens.
XX
PN      US2003104529-A1.
XX
PD      05-JUN-2003.
XX
PF      04-JAN-2002; 2002US-00037270.
XX
PR      21-JAN-2000; 2000US-00488725.
PR      25-APR-2000; 2000US-00552317.
PR      19-JUL-2000; 2000US-00620312.
XX
PA      (ZHOU/) ZHOU P.
PA      (TANG/) TANG Y T.
PA      (LIUC/) LIU C.
PA      (ASUN/) ASUNDI V.
PA      (DRMA/) DRMANAC R T.
PI      Zhou P, Tang YF, Liu C, Asundi V, Drmanac RT;
PI      WPI; 2003-678194/64.
XX
PT      New polynucleotide, useful for treating diseases e.g., cancer or
PT      neurodegenerative diseases.
XX
PS      Claim 1; SEQ ID NO 768; 99pp; English.
XX
CC      The invention relates to a polynucleotide comprising a sequence given in
CC      the specification, or its mature protein-coding portion, or its
CC      complement. The polynucleotide is useful for treating diseases e.g.,
CC      cancer or neurodegenerative diseases and many others listed in the
CC      specification. The present sequence represents a novel human cDNA. Note:
CC      The sequence data for this patent did not form part of the printed
CC      specification but was obtained in electronic format directly from USPTO
CC      at Seqdata.uspto.gov/sequence.html?docid=20030104529.
XX
SQ      Sequence 1844 BP; 535 A; 380 C; 389 G; 540 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 1,42e-184 Length: 1844
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: Gaps: 1
US-10-077-111-13 (1-384) x ADB48858 (1-1844)
Qy      1 MetValLysLeu1IleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe 20
Db      164 ATGGTGAATTAAGTATTCACATTAATAGTGAATGATGATGATGATGATGATGATGATGATG 223
Qy      21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThr1IleArgLeuTrpSerLeuArg 40
Db      224 TCCCTTTCCTCTCTGCTACTCTCTCTCTGACAAACAAATTCGCTGATCTGTTAAGT 283
Qy      41 AspPheThrGlnLeuProHisSerProLeuLysPheHisThrTrpAlaValHisCysCys 60
Db      284 GACTTATCTAAGTACGACATTCCTCAATTAAGTTTCAATCAATGCTGCTGCTGCTGCTG 343
Qy      61 CysPheSerProSerGlyHis1IleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
Db      344 TGTTCCTCCCTTCAGGACATATTTTGGCATGCTGTTCACACATGATGATGATGATGATG 403

```

RESULT 11
ADB48858

QY 81 TTPAsnThrGluanglgiMleuAlaValMeGluGlnProSerGlySerProVal 100
 Db 404 TGGAAATACGAAATAGGACAGATGCTGCGAGTGTGAAACAGGCTAGTGGCAGCCCTGTG 463
 QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
 Db 464 AGGGTTTGCAGTTTCCAGACTCCACGTTGTCATCAGGGGCGAGCTGATGAACT 523
 QY 121 ValValLeuThrPheAlaGlnSerTyrLysLeuThrArgCysGlySerValLysAspGly 140
 Db 524 GTGGTTTGTGAAATGACAGCTCATCAAAATTATATGATGTGTTAAAGATGGC 563
 QY 141 SerLeuAlaAlaCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCys 160
 Db 584 TCCCTGGCGGCGAGTGCATTTTCTCTAATGGAAGCTCTTTGTCACTGGCTCCTCATGT 643
 QY 161 GlyAspLeuThrValTTPAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
 Db 644 GGTGATTTAAACAGTGTGGATGATAAATAGAGTGTCTGCATATGTGAAAAAGCACATGAT 703
 QY 181 LeuGlyLysLeuThrCysGlyAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
 Db 704 CTGGAAATTACCTGCTGCGATTTTCTTCAAGCCAGCTTCTGATGGAGAACAGGCTCT 763
 QY 201 GlnPhePheArgLeuAlaSerCysGlyLysAspCysGlnValLysIleTyrPileValSer 220
 Db 764 CAGTTTTTGCAGCTGGCATATGTCAGATTTGCCAAGTCAAAATTGGATTGTTCT 823
 QY 221 PheThrHisIleLeu----- 225
 Db 824 TTAAACCATATCTTAGGTTTGAATTAATAATATAAAAGTACACTGAGTGGCACTGTGCT 883
 QY 225 ----- 225
 Db 884 CCGTCTCTGCGCTGTGCTTTTCCCATGATGGCAGATGCTAGTCTCAGGGTCACTGGAT 943
 QY 225 ----- 225
 Db 944 AAGCTGTCTAATGATATGATACTAATACAGAGAAATTAATCTTCAACATGACTCAGCAC 1003
 QY 225 ----- 225
 Db 1004 ACCAGTATGTCAACACTGTGCTTTGACACTAATACCTTTTACTTGCTACTGTGTCA 1063
 QY 226 -----AlaArgArg 228
 Db 1064 ATGGACAAAACAGTGAACATCTGGCAATTGACCTGGAACACTTTGCCAAGCAAGAGC 1123
 QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTyrSerGluGluValValSerThrTyr 248
 Db 1124 ACAGAACATCAGCTGAGAGCAATTTTACCGAAGATTGGTCAGAGAGAGATGTCTCAACATCG 1183
 QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyLysPheLysMetAsnAsnIleAspGly 268
 Db 1184 CTGTGTGCACAAAGTTTAAAGATCTGTGTGATTTTCAAGATGAATTAACATTGATGGA 1243
 QY 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
 Db 1244 AAAGAACCTGTGATCTTCAAAAAGAAAGCTGCTCATATTTGAAAATTTGATATCTCTA 1303
 QY 289 GlyLeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeu 308
 Db 1304 GGACTGTGGTAAAGTCTGAGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTT 1363
 QY 309 SerSerGlyLysProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
 Db 1364 TCTTCAGGAATCTCTGATGAATTTATATGTCATTAATCTAGAGAACTTATGAAAAGATCCG 1423
 QY 329 ValIleAlaSerAspGlyTyrSerTyrGluLysGluAlaMetGluAsnTyrPileSerLys 348
 Db 1424 GTCATCCGATCAGATGCTATTTCAATGAAAAAGAACCAATGCAAAATTTGATCAGCAAA 1483

QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
 Db 1484 AAGAAACGTACAAAGTCCATGACAAATCTTGTCTTCCCTCAGCGGTACTTACCAACAAAT 1543
 QY 369 ArgThrLeuLysMetAlaIleAsnArgTyrPheGluThrHisGlnLys 384
 Db 1544 AGGACTCTGAAAATGGCCATCAATGATGGCTGGAGACACACAAAG 1591
 RESULT 12
 ID AAD45076 standard; DNA; 1908 BP.
 AC AAD45076;
 DT 27-DEC-2002 (first entry)
 XX
 DE Human RET16.3 splice variant DNA.
 KW Human; RET16; intracellular signal; inflammation-related disease; asthma; rheumatoid arthritis; psoriasis; multiple sclerosis; neuroprotective; KW transplacental rejection; chronic obstructive pulmonary disease; TNF-alpha; KW inflammatory bowel disease; Crohn's disease; immunosuppressive; cancer; KW inacute respiratory distress syndrome; cardiac; ulcerative colitis; KW autoimmune disease; cystic fibrosis; gene therapy; cytostatic; neoplasm; KW cellular migration disorder; cell proliferation disorder; calcification; KW hyperinsulinaemia; diabetes type 2; systemic lupus erythematosus; tumour; KW cardiovascular disease; Wegener's granulomatosis; atherosclerosis; KW thalassemia; vasotropic; gene; ds.
 XX Homo sapiens.
 OS
 FH Key Location/Qualifiers
 FT 136..164
 FT CDS /tag= a
 FT /product= "Human RET16.3 splice variant protein"
 WT MO20026494-A2.
 PD 29-AUG-2002.
 PF 15-FEB-2002; 2002MO-US005162.
 XX
 PR 16-FEB-2001; 2001US-0269366P.
 PR 29-MAY-2001; 2001US-0294181P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PT Todderud CG, Finger JN, Rillema J;
 DR WPI; 2002-682760/73.
 XX P-PSDB; AAE28168.
 XX
 PT New human, mouse or rat RET16 genes and proteins, involved in
 PT intracellular signaling cascade, useful for in gene therapy, particularly
 PT for treating e.g. ischemia, cystic fibrosis, autoimmune disease, cancers,
 PT tumors or neoplasms.
 PT
 XX
 PS Claim 1, Page 168-169, 175pp; English.
 XX
 CC The invention relates to human, mouse or rat RET16 genes and proteins,
 CC involved in intracellular signaling cascade. The RET16 protein or
 CC polynucleotide is useful for treating an inflammation-related disease or
 CC disorder, e.g. rheumatoid arthritis, juvenile arthritis, psoriasis, and
 CC asthma, ischaemia-reperfusion, multiple sclerosis, rejection of organ or
 CC tissue transplants, chronic obstructive pulmonary disease, inflammatory
 CC bowel disease, Crohn's disease, ulcerative colitis, inacute respiratory
 CC distress syndrome, systemic lupus erythematosus, autoimmune disease,
 CC cystic fibrosis, cancers, tumours or neoplasms. This inflammation-related
 CC disease or disorder also includes disorders associated with aberrant
 CC activation of the TNF-alpha pathway, disorders associated with aberrant
 CC cellular migration, proliferation, metastasis, juvenile idiopathic
 CC arthritis, haematogenous metastases of tumour cells, hyperinsulinaemia,
 CC diabetes type 2, atherosclerosis, cardiovascular disease, colon cancer,

PI Mao Y, Xie Y,
 XX WPI; 2001-537038/60.
 DR P-PsDB; AAG78660.
 XX
 PT New polypeptide-beta-transducin 41 for treating e.g. cancer and HIV
 PT infection.
 XX
 PS Claim 6; Page 24-25 (Disclosure); 32pp; Chinese.
 XX
 CC The present invention provides the protein and coding sequences of beta-
 CC transducin 41. The sequences can be used in the treatment of cancer and
 CC HIV infection. The present sequence is the coding sequence of the
 CC invention
 XX
 SQ Sequence 1623 BP; 485 A; 302 C; 335 G; 501 T; 0 U; 0 Other;

Alignment Scores:

| Pred. No.: | 1,966-172 | Length: | 1623 |
|------------------------|-----------|---------------|------|
| Score: | 1854.00 | Matches: | 364 |
| Percent Similarity: | 79.18% | Conservative: | 1 |
| Best Local Similarity: | 78.96% | Mismatches: | 2 |
| Query Match: | 90.57% | Indels: | 94 |
| DB: | 4 | Gaps: | 1 |

US-10-077-111-13 (1-384) x AA164914 (1-1623)

```

QY 18 CysAlaPheSerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyr 37
DB 3 TGTGCTCTTCTCTTTCTCTCTCTGCTACTTCTCTCTTGGACAAACAAATTCGCTGTAC 62
QY 38 SerLeuArgAspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaVal 57
DB 63 TCGTTAGTACTTACTTACTGAACTGCACATTCCTCCATGAAATTACTTACTTACTGCTC 122
QY 58 HisCysCysCysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThr 77
DB 123 CACTGCGCTGTTCTCTCTCTGAGCATAATTTGGCATGCTGTTCAACAGATGGTACC 182
QY 78 ThrValLeuThrAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGly 97
DB 183 ACTGCTCTATGAAATCTGAAATGACAGATGCTGAGATGAGAAACCTTACTGAGC 242
QY 98 SerProValArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAla 117
DB 243 AGCCCTGTGAGGCTTCCAGTTTCCCAAGCTCCACGCTTTGGCATCAGGGCAGCT 302
QY 118 AspGlyThrValValLeuThrAsnAlaGlnSerTyrIleLeuTyrArgCysGlySerVal 137
DB 303 GATGGAACCTGCTGTTGTGAAATGACAGTCAATACAAATTAATAGATGGTACTGTT 362
QY 138 LysAspGlySerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGly 157
DB 363 AAAGATGGCTCTTGGGCGCATGTCATTTCTCTTAATGAAAGCTTCTTGTACAGGC 422
QY 158 SerSerCysGlyAspLeuThrValThrAspAspLysMetArgCysLeuHisSerGlyLys 177
DB 423 TCCCTCATGTGTGATTTAACTGCTGAGATGATAAATGAGTGTCTGCTAGTGAATAA 482
QY 178 AlaHisAspLeuGlyIleThrCysAspPheSerSerGlnProValSerAspGlyGlu 197
DB 483 GCACATGATCTTGGAAATTAATCTGCTGCAATTTTCTTCAACAGCCAGTTTCTGATGAGAA 542
QY 198 GlnGlyLeuGlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTyr 217
DB 543 CAAGGCTTCAAGTTTTCAGCTGGCATATGTGGTACAGATTTGCCAAAGTCAAAATTTGG 602
QY 218 IleValSerPheThrHisIleLeu----- 225
DB 603 ATTGTTTCTTTTACCCTATCTTACTGATTGATTAATAATAAAGTACAGTCAAGTGG 662
QY 225 ----- 225

```

```

DB 663 CACTGTCTCTGTTCTGTGCTGCTGCTTTTCCCATGATGGCAGATGCTACTTCAGGG 722
QY 225 ----- 225
DB 723 TCAGTGAATAGCTGTCTCATATGATATGATATGATATGATATGATATGATATGATATG 782
QY 225 ----- 225
DB 783 ACTCAGACACACAGATATGTCACAACTGTGCTTTTGCACATTAATACCTTTTACTTGTCT 842
QY 225 ----- 225
DB 843 ACTGTTCAATGACACAAACAGTGAACATCTGCAATTTGACCTGGAACACTTGTCCAA 902
QY 226 AlaArgArgThrGlnHisGlnLeuLysGlnPheThrGluAspTyrSerGlnGluValVal 245
DB 903 GCAGGCGGACAGAACATACGCTGAAGCAATTTACGAAAGATTGGTCAAGAGAGATGTC 962
QY 246 SerThrTyrLeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsn 265
DB 963 TCACATGCGCTTTGTGCACACAGATTTAAAGATCTTGTGTATTTTCAAGATGAATAAC 1022
QY 266 IleAspGlyLysGlnLeuLeuLeuLeuThrLysGlnSerLeuAlaAspAspLeuLysIle 285
DB 1023 ATTGATGAAAGAAAGACTGTGAATCTTACAAAGAAAGCTGCTGATGATTTGAAATTT 1082
QY 286 GluSerLeuGlyLeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLysVal 305
DB 1083 GAATCTCTAGACTGTGTATGTAAGTCTGAGAGAAATTTGAAGAGCTCAGACCAAGTT 1142
QY 306 LysSerLeuSerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMet 325
DB 1143 AAATCCCTTTCTCAGGAATTCCTGATGAATTAATATGTCATTAACCTGAGAACTTATG 1202
QY 326 LysAspProValIleLeuAspArgLysTyrSerTyrGluLysGlnAlaMetGluLeuTyr 345
DB 1203 AAAGATCCGCTCATGTCATCAATGAGTCTATTCATATGAAAGAGAAAGCAATGAAATTTGG 1262
QY 346 IleSerLysLysLysArg-ThrSerProMetThrAsnLeuValLeuProSerAlaValLe 365
DB 1263 ATCAGCAAAAGAAAGAGGATCAAGTCCCATGCAAAATCTTGTCTTCAAGCGGCT 1322
QY 365 u-ThrProAsnArgThrLeuLysMetAlaIleAsnArgTyrLeuGlnThrHisGlnLys 384
DB 1323 TTACACCAATATGAACTTGAAATGCGCATCATATGATGCTGAGACACACCAAAAG 1381

```

RESULT 14
 AA160662/c
 ID AA160662 standard; cDNA; 1826 BP.
 XX
 AC AA160662;
 XX
 DT 22-OCT-2001 (first entry)
 XX
 DE Human polynucleotide SEQ ID NO 4651.
 XX
 KW Human; neotropic; immunosuppressant; cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's; Parkinson's disease; Huntington's disease; hemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemoclastic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukemia; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153312-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 26-DEC-2000; 2000WO-US034263.
 XX
 PR 23-DEC-1999; 90US-00471275.
 PR 21-JAN-2000; 2000US-00488725.

OY 384 s 384
Db 262 A 262

RESULT 15
ADM19385
ID ADM19385 standard, cDNA, 1686 BP.

XX
AC ADM19385;
XX
DT 20-MAY-2004 (first entry)
XX
DE Novel human channel/transporter gene #182.
XX
da; gene, immunosuppressive; antiarthritic; antiinematic;
KM antiproliferative; cytosolic; cardiac; vasoprotective;
KM nocitropic; neuroprotective; antibacterial; virucide; fungicide;
KM opthalmological; gene therapy; channel/transporter protein;
KM rheumatoid arthritis; neoplasm; cardiac arrest; cerebrovascular disorder;
KM cerebral ischemia; angiogenesis; nervous system disorder;
KM Alzheimer's disease; ocular disorder; corneal infection; wound healing;
KM epithelial cell proliferation; skin aging; sunburn; transplantation;
KM chemotaxis; food additive.

XX
OS Homo sapiens.
XX
PN WO200154472-A2.
XX
PD 02-AUG-2001.
XX
PF 17-JAN-2001; 2001WO-US001307.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215115P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218230P.
PR 26-JUL-2000; 2000US-0220963P.
PR 14-AUG-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226682P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235844P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240609P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.


```

QY      289 GlyLeuArgSerLyValLeuArgLysIleGluGluLeuArgThrLySValLySerLeu 308
Db      1280 ----- 1280
QY      309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
Db      1280 ----- 1280
QY      329 ValIleAlaSerAspGlyTyrSerTyrGluLyGluAlaMetGluAsnTrpIleSerLyS 348
Db      1281 -----GATGGCTATTTCATATGAAAAGAGAGCAATGAGAAATTGGATCAGCAAA 1328
QY      349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db      1329 AAGAAACGTACAGTCCCATGACAAATCTGTCTTCCTTCACGGGTACTTACACCAAAAT 1388
QY      369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys 384
Db      1389 AGGACTCTGAAATGGCCATCAATAGTGGCTGGAGACACCAAAAG 1436

```

Search completed: February 5, 2005, 12:07:55
 Job time : 723 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 11:39:07 ; Search time 4204 Seconds
(without alignments)
3476.850 Million cell updates/sec

Title: US-10-077-111-13
Perfect score: 2047
Sequence: 1 MVXLIHTLADHGDDVNCACF.....LTPNRLKMAINWLETHOK 384

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Xgapop 10.0 , Xgapext 0.5
Xgapop 6.0 , Xgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 segs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=rlh
-O=/cgn2_1/USFTO.spool/US10077111/runat_04022005_121053_6816/app.query.fasta.1.583
-DB=EST -QMT=faetap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=400 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US10077111 @CGN 1.1 4352 @runat_04022005_121053_6816 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEDBTRY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MAIN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST: *
1: gb_est1: *
2: gb_est2: *
3: gb_hic: *
4: gb_est3: *
5: gb_est4: *
6: gb_est5: *
7: gb_est6: *
8: gb_ges1: *
9: gb_ges2: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|----------------------|
| 1 | 1977 | 96.6 | 1770 | 3 | CR607851 full-length |
| 2 | 1714.5 | 83.6 | 1600 | 3 | CR619104 full-length |
| 3 | 1617 | 79.0 | 2027 | 3 | AK011391 Mus muscu |
| 4 | 1205 | 58.9 | 889 | 1 | AL553333 AL553333 |
| 5 | 1203 | 58.8 | 941 | 1 | AL533510 AL533510 |
| 6 | 1154 | 56.4 | 1010 | 4 | BM543484 AGENCOURT |
| 7 | 1043.5 | 50.4 | 858 | 4 | BI603184 603249828 |
| 8 | 1031.5 | 50.4 | 792 | 4 | BI821143 603035030 |
| 9 | 1005 | 49.1 | 728 | 7 | CN793033 4127990 B |

| | | | | | |
|----|-------|------|------|---|--------------------|
| 10 | 975.5 | 47.7 | 910 | 5 | B0895162 |
| 11 | 943.5 | 46.1 | 838 | 4 | BI107478 |
| 12 | 907 | 44.3 | 839 | 7 | CK306752 SB02035A1 |
| 13 | 883.5 | 43.2 | 808 | 7 | CK792250 AGENCOURT |
| 14 | 880 | 43.0 | 499 | 5 | BX282729 BX282729 |
| 15 | 880 | 43.0 | 636 | 4 | BI759505 |
| 16 | 879 | 42.9 | 727 | 7 | CK364577 |
| 17 | 873 | 42.6 | 505 | 1 | AI189142 |
| 18 | 872 | 42.6 | 807 | 1 | AL040518 |
| 19 | 856.5 | 41.8 | 817 | 5 | B0961462 |
| 20 | 854 | 41.7 | 718 | 5 | B0356767 |
| 21 | 834 | 40.7 | 652 | 2 | BB653743 |
| 22 | 821 | 40.1 | 669 | 1 | AJ726222 |
| 23 | 801 | 39.1 | 869 | 5 | BU148588 |
| 24 | 797 | 38.9 | 760 | 5 | BU273077 |
| 25 | 758.5 | 37.0 | 965 | 4 | BG769613 |
| 26 | 758 | 37.0 | 683 | 6 | CD215933 |
| 27 | 747 | 36.5 | 919 | 5 | BQ713478 |
| 28 | 746 | 36.4 | 791 | 7 | CK603751 |
| 29 | 739 | 36.1 | 682 | 6 | BY710675 |
| 30 | 725 | 35.4 | 768 | 7 | CK846814 |
| 31 | 719 | 35.1 | 559 | 4 | BI681850 |
| 32 | 715 | 34.9 | 967 | 1 | AL533462 |
| 33 | 705 | 34.4 | 797 | 5 | BU423666 |
| 34 | 700 | 34.2 | 764 | 4 | BG619844 |
| 35 | 698.5 | 34.1 | 767 | 7 | CNS26819 |
| 36 | 691 | 33.8 | 723 | 9 | CE487278 |
| 37 | 690 | 33.7 | 632 | 4 | BU624489 |
| 38 | 677 | 33.1 | 626 | 4 | BG562985 |
| 39 | 673.5 | 32.9 | 447 | 1 | AA118718 |
| 40 | 668.5 | 32.7 | 1042 | 4 | CO567478 |
| 41 | 662.5 | 32.4 | 829 | 7 | CN407189 |
| 42 | 662 | 32.3 | 463 | 7 | CN407189 |
| 43 | 662 | 32.3 | 897 | 5 | BQ964199 |
| 44 | 657 | 32.1 | 1010 | 3 | BC039804 |
| 45 | 648 | 31.7 | 699 | 1 | AJ726210 |

ALIGNMENTS

RESULT 1
CR607851
LOCUS full-length cDNA clone CS0DN0047015 of Adult brain of Homo sapiens
DEFINITION (human).
ACCESSION CR607851.1 GI:50488658
VERSION CR607851.1
KEYWORDS HTC; CNS/IT cdna.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1770)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
REMARK Contact: Feng Liang Email: fliang@life.technet.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600
Faraday Avenue
Genoscope.
2 (bases 1 to 1770)
REFERENCE Direct Submission
TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web: www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1..1770
/organism="Homo sapiens"

Db 823 ATTACCGGTCTTAGGCTTGAATTAATAATTAAGACACATGAGTGGGACTGGCC 882
Qy 222 ----- 222
Db 883 CCTGTTCGGCTGTGCTTTTTCATCATGATGGAAGAATGCTTGCACTGGGGTCACTGAT 942
Qy 223 ----- 223
Db 943 AAATCTGTATCATATACATGATGATGGCCCTCAGAGTGTGTACACAGCTGATCAGAT 1002
Qy 229 ThrluHs-----HsileuHsAarg 228
Db 1003 ACCAGGATGTATGACGATTTGGCTTGACCCCAACTCTTACTTACTGATGTTCA 1062
Qy 231 ----- 231
Db 1063 ATGACAAAGACAGTGAACATTTGGCAGTTTGACTGGAAACACTTGGCAAGAACATG 1122
Qy 232 -----GluLeuYsgInPheThrGluAspTrpSerGluGluValSerThrTrpLeu 249
Db 1123 AACGACCCGCTGAAACATTTTCACTGAGAGATGTCAGAGAGATGTCCTCGTGTGGTT 1182
Qy 250 CysAlaGlnAspLeuValGlyLeuPheLeuMetAsnGlnAspGlyLys 269
Db 1183 CGTCTCAAGGCTTGGAAAGACCTGCTGATTTTTCAGGGCAACATCATGATGGGAA 1242
Qy 270 GluLeuLeuAsnLeuThrLysGluSerLeuAlaAspLeuValGluSerLeuGly 289
Db 1243 GAATCTATTGCACTTCAACAAAGAAAGTGTGGCTGTGATTTGAAATGAAATCTCTAGGG 1302
Qy 290 LeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeuSer 309
Db 1303 CTGGCAGCAAGGCTGAGAGATTTGAAGAGCTCAGGGCCAGATGATGATCTCTCT 1362
Qy 310 SerGlyLeuProAspGluPheIleCysProIleThrArgGluLeuMetLysAspProVal 329
Db 1363 TCCGGAATCCCTGACGATTCATCTGCCCAATTAACAGAACTCATGAAGAACCCGCTC 1422
Qy 330 IleLeuSerAspGlyLysTrpSerLysGluLeuValMetGluLeuValSerLysLys 349
Db 1423 ATCCATCAAGTGGCTACTCTTACAGAGAACCAATGAAACTGTGATCACAAAGAG 1482
Qy 350 LysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsnArg 369
Db 1483 AAGGTACGAGCCCAATGACAAATTTGGCTCTCCCTTCACTGATGATGACCCCAAGAG 1542
Qy 370 ThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys 384
Db 1543 AACTGAAGATGGCCATCAACCGATGGCTGAGAGCGACAGAGAG 1587
RESULT 4
AL553333 889 bp mRNA linear EST 30-MAR-2004
LOCUS AL553333 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS001075YB08 5-PRIME, mRNA sequence.
ACCESSION AL553333
VERSION AL553333.3 GI:45858102
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 889)
AUTHORS L.J.W.B., Gruber,C., Jeesee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31275147.
Contact: Genoscope - Centre National de Sequencage
Genoscope - Centre National de Sequencage
2 rue Gaëton Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned

into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1653.r
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna/cds/CS001075D040P1&c=1653.r>.
FEATURES
source
1..889
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS001075YB08"
/rname="PLACENTA COT 25-NORMALIZED"
/rsize="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Alignment Scores:
Pred. No.: 5,48e-112 Length: 889
Score: 1205.00 Matches: 222
Percent Similarity: 99.12% Conservative: 2
Best Local Similarity: 98.23% Mismatches: 1
Query Match: 58.87% Indels: 1
DB: 1 Gaps: 0
US-10-077-111-13 (1-384) x AL553333 (1-889)
Qy 1 MetValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCysCysAlaPhe 20
Db 112 ATGTGTGAACATGATTCACACATTAAGTCGATGATGATGATGATGATGATGATGATGAT 171
Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuYrSerLeuArg 40
Db 172 TCTTTTCCCTTGTGGCTACTGTGCTTGTGACAAACATTCGCTGATCGTATCGT 231
Qy 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrIleAlaValHisCysCys 60
Db 232 GACTTAACTGAACATGACACATTTCTCCATTAAGTTTCATGCTGATGCTGCTGCTGCTG 291
Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
Db 292 TGTTTCTCCCTTCAGACATATTTTGGCATCGTTTCAACAGATGATGATGATGATGATGATG 351
Qy 81 TrpAsnThrGluLysGlnGlyMetLeuAlaValMetGluLysProSerGlySerProVal 100
Db 352 TGGAAATGATGAATATGACAGATGCTGCGATGATGATGATGATGATGATGATGATGATG 411
Qy 101 ArgValCysGln-PheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyTh 120
Db 412 AGGTTTGGCCAGTTTTCCTCCAGACATCCAGTGTGTTGGCATGAGGGGCACTGATGAGAC 471
Qy 120 ValAlaValLeuThrAsnAlaGlnSerLysLysLeuYrArgCysGlySerValLysAspG 140
Db 472 TGTGATTTTGTGAAGTACAGATCATCAATATTAATATGATGATGATGATGATGATGATG 531
Qy 140 YSerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCy 160
Db 532 CTCCTTGGGGGAGATGTCCTTTTCTCTATGAGAGTTCTTTGTGTCATGCTGCTCATG 591
Qy 160 GGIAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGluLysAlaHisAs 180
Db 592 TGGGATTTTATCAGTGGGATGATTAAGAAGAGTGTCTGCATATGTAAGAAACACATGA 651
Qy 180 PLeuGlyIleThrCysGlyAspPheSerSerGlnProValSerAspGlyGluGlnGlyLe 200
Db 652 TCTTGAAATTAACCTGCTGCGATTTTCTTACAGCCAGTTTCATGATGAGAAACAGAGTCT 711
Qy 200 uGlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSe 220
Db 712 TCAGTTTTCGACTGCAATCATGTGTGATGATGATGATGATGATGATGATGATGATGATG 771

| | | | |
|----|-----|---|-----|
| Oy | 101 | ArgValaCygSgInPheSerProaSerPheThrCysLeuAlaIaSerGlyValAlaIaAspGlyThr | 120 |
| Db | 451 | AGGGTTTCCCAAGTTTTCCCCAGACTCCACGCTTTGGCATCAAGGGGCAAGCTGATGAACCT | 510 |
| Oy | 121 | ValValIleuTPaSmAlaGInSerTybYsLeuTYaGcYsGlySerValIlyAspGly | 140 |
| Db | 511 | GTGGTTTTGTGGAAAGCAGATCATACAAATTATATAGATGTGGTAAGTTAAAGATGGC | 570 |
| Oy | 141 | SerLeuAlaAlaCyValaPheSerProaAngIySerPhePheValThGlySerSerCyv | 160 |
| Db | 571 | TCCTTGGCGGCGATGGCAATTTCTCTCTAATGGAAAGCTTTGTGCATCGCTCCATCATGT | 630 |
| Oy | 161 | GlyAspLeuThValTPaAspAspYsMetAngCyLeuThIaSerGlyValAlaIlaAs | 180 |
| Db | 631 | GGTGATTTTAAACAGTGTGGATGATCAATCAAGAGGGTGTGTGCATATGTGAAGAAAGCACAATGG | 690 |
| Oy | 180 | PleuGlyIleThrCysCyAsp---PheSerSerGln-ProValSerAsp-GlyGluGln | 198 |
| Db | 691 | ATCTGGAATTACCTGCTGCGCATTTTTTTCTTCACAAAGCAAGTCTCTGATTGGAAACAA | 750 |
| Oy | 199 | GlyLeuGln--PhePheArgLeuAlaSer-CysGly-GlnAspCyGlnValIlyIleThr | 217 |
| Db | 751 | GATCTTCAAGTTTTTTTCGAAGTGCATCGCATGTGGCCAGGATTCGCAAGTCCAAATATG | 810 |
| Oy | 217 | PleValSer-PheThrHisIleLeuAlaIaArgAThGlu | 230 |
| Db | 811 | GGATGGGTCCTTTTACCGATATCTTGCAAGGGGACACAG | 851 |

| LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL | COMMENT |
|----------|--|-----------|------------|-------------|--------|----------------------|--------------------|--|--------------------|---|--|
| BI821143 | 60303503071 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176086 5', mRNA sequence. | BI821143 | BI821143.1 | GI:15932693 | EST. | Homo sapiens (human) | 1 (bases 1 to 792) | National Institutes of Health, Mammalian Gene Collection (MGC) | Unpublished (1999) | Contact: Robert Strausberg, Ph.D. Email: c9apbs-r@mail.nih.gov | Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLML1438 row: 1 column: 07 High quality sequence etop: 751. |

FEATURES

```

1. .792
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5176086"
/lab_host="DH10B"
/clone_1lb="NH MGC 115"
/note="Organ: pooled brain, lung, testis; Vector:
pcmv-sport; site_1: NotI; site_2: EcoRV (deinserted); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C
Gruber (Invitrogen). Research Genetics tracking code
221. Note: this is a NH MGC Library."

```

| ORIGIN | |
|------------------------|----------|
| Alignment Scores: | |
| Pred. No.: | 2-27e-94 |
| Score: | 1031.50 |
| Percent Similarity: | 82.21% |
| Best Local Similarity: | 91.77% |
| Query Match: | 50.39% |
| DB: | 4 |
| | |
| Length: | 792 |
| Matches: | 212 |
| Conservative: | 1 |
| Mismatches: | 10 |
| Indels: | 9 |
| Gaps: | 1 |

US-10-077-111-13 (1-384) x BI821143 (1-792)

QY 1 MetValylsleuIleHisThrLeuAlaAspHisGlyAspAspValAsnCysGlyAlaPhe 20
Db 118 ATGGGAACATGATTACACATTAAGTCTATCATGTGTGACGATGTCACTGGTGGCTTTC 177
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40
Db 178 TCCTTTTCCTCTGGCTACTGCTC--CTGGACAAACAAATTCCGCTGATCTGTTACGT 234
QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys 60
Db 235 GACCTTATCTGAATCTGCACATCTCCATTGAAATTTCATTACTTAAGCTGTCCACTGCTGC 294
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
Db 295 TGTTCCTCCCTTACGACATAT-TTGGATCGTGTTCACAGATGATACCACTGCTCCA 353

| | | | |
|----|-----|---|-----|
| QY | 81 | TPPAAETPRNGIUAENGILGIMETLEUAAVALMETGLUGINPROSERGIYSEPRVAL | 100 |
| Db | 354 | TGGAACTACTGAAAATGACACAGTCTGGCAGATGGAAACAGCTTATGGCAGCCTGTG | 413 |
| QY | 101 | ARGVALCYGGINPHESEPRCHAPSERTHRQYLEUAAJASERGILYALAAAPGIYTHR | 120 |
| Db | 414 | AGGGTTTGCCAGTTTCCCCAGACTCCACAGTGTGGCATCAGGGGCACTGATGGAACT | 473 |
| QY | 121 | VALVALLEUTPRAAHLAGINSERYTYLYSLEUTYRANGCYGGLYSERVALLYSAPRGY | 140 |
| Db | 474 | GTGGTTTGTGGAAAGCACAGTCATCAACAAATTATAGTGGTGTAAAGATGGAC | 533 |
| QY | 141 | SERLEUAAJALACYALAPHESERPROANGIYSERPHEPHEVALTHFGIYSESER-CY | 160 |
| Db | 534 | TCCTTGGGGGAGATGGCACTTTCTCTTAATGGAACCTCTTTGTCACTGGCTCCTAAG | 593 |
| QY | 160 | EGIYAPLEUTHRVALTPAAPAPPLYMECARQCYLYEUIHISERGILU-LYVALAHIAA | 180 |
| Db | 594 | TGGTATTATTAACAGGTGGAGATGAATAAATGAGGTCTGCATAGTAATTAAGCCCAAG | 653 |
| QY | 180 | APLEUGILILETHR-CYSCYAPHPHESERSEGINPROVALSER-APPGIYGLUGINGI | 199 |
| Db | 654 | ATCTTGGAAATTAACCTTGCTGCAGATTTTCTTACAGCCAGTTTCTTGATGGAGAACAGG | 713 |
| QY | 199 | YLEUGINPHEPHEARG-LEUAAJASERCYS-GIYGINAPCYGGLINVALLYS-ILETRPII | 218 |
| Db | 714 | TCTTCAGTTTCTCGACTGGCATCATGGGCTCAGAGATTGGCATGTCAACAATTTGGAC | 773 |
| QY | 218 | EVALSERPHEPHEHISILE 224 | |
| Db | 774 | TGTNACTTTTACCACATATT 792 | |

| | | | | | |
|------------|---|-------------|------|----------|----------------|
| RESULT 9 | CN793033 | 728 bp | mRNA | 1 linear | EST 26-MAY-200 |
| LOCUS | CN793033 | | | | |
| DEFINITION | 4127990 BARC 8BOV Bos taurus CDNA clone 8BOV_46H02 5', mRNA sequence. | | | | |
| ACCESSION | CN793033 | | | | |
| VERSION | CN793033.1 | GI:47689013 | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Bos taurus (cow) | | | | |
| ORGANISM | Bos taurus | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos. | | | | |

| | | | | |
|------------|---|--------|---|------------------------|
| OY | | 1 | MctValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValAsnCyCysValAlaPhe | 20 |
| Db | | 5 | ATGGTGAAGCGCTGATTCC-ACGCTGGGTATTCACGGTGATGACTGACGCTGCCTGCCCTTC | 63 |
| OY | | 21 | SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyzSerLeuArg | 40 |
| Db | | 64 | TGGGCTGGCCCTCTGGCCACCTGCTCTTGGACAAAGCATCCGGCTGTAACCTTAAGT | 123 |
| OY | | 41 | AspPheThrGluLeuProHissereProleuLysPheHisThrTYrAlaValHisCyCys | 60 |
| Db | | 124 | GACTTTGCTGAACCTCCGTAACCCCCCGCGAAGTTCCACACTAAGCTGTCCAACGTCTGC | 183 |
| OY | | 61 | CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrrThrValLeu | 80 |
| Db | | 184 | TGTTTCTCACTCCACAGACAGCTTTTAGCATGTGTGTGCACAGAGGAGCACCGGTGCTG | 243 |
| OY | | 81 | TrpAenThrGluAsnGlyGlnmetreLeuAlaValmetGluGlnProserGlySerProVal | 100 |
| Db | | 244 | TGGAGACTGGACAGAGGGAGCACACCCTGAACCGTGTGGAGACACCGGGTGGACGCCCTGG | 303 |
| OY | | 101 | ArgValCYsGlnPheSerProAspSerThrCYsLeuAlaSerGlyAlaAlaAspGlyThr | 120 |
| Db | | 304 | CGCGTCTGTGCTTTTCCCACAGACTCTGCTTACTTAAGCGTCAGAGGGCTGCCGATGATCC | 363 |
| OY | | 121 | ValValLeuThrTrpAsnAlaGlnserTyzLysLeuTYrArgCYsGlySerValLysAspGly | 140 |
| Db | | 364 | ATGCTGTTGTGGAAATGACACAGACATPAACAATATAGGTGTGTGATCAAGAATAGC | 423 |
| OY | | 141 | SerLeuAlaAlaCYsAlaPheSerProAsngLySerePheValThrGlySerSerCys | 160 |
| Db | | 424 | TCATTGGTGGGCTGTGGCGTTTTCTCCCAATGAGAGGCTTTGTCTACGTGCTCTCGGGC | 483 |
| OY | | 161 | GlyAspLeuThrValITrpAspAspLysMetArgCYsLeuHisSerGlyLysAlaHisAsp | 180 |
| Db | | 484 | GGGGACTTGACAGTGTGGGATGACAGAAATGAGGTGTCTACACAGCAGAGAGGCGCACGAT | 543 |
| OY | | 181 | LeuGlyIleThrCYsCYsAspPheSerSerGlnProValSerAspGlyGluGlnIleu | 200 |
| Db | | 544 | CTGGGATATCACTGTGTGACGCTTTCTCTCACAGCCTCTCTCTGGGGAGAA---GGCCTC | 600 |
| OY | | 201 | GlnPhePheArgLeuAlaSerCYsGlyGlnAspCYsGlnValIleIleThrIleValSer | 220 |
| Db | | 601 | CAGTCTTTCACAGTTGGCGTCATGTGTGTCAAGCTGTGAATCAAACCTCTGGGCTTTACT | 660 |
| OY | | 221 | PheThrHisIleLeu 225 | |
| Db | | 661 | ATTACCCCGTCTTA 675 | |
| <hr/> | | | | |
| RESULT 11 | | | | |
| BII07478 | | | | |
| LOCUS | BII07478 | 838 bp | mRNA | linear EST 26-JUN-2001 |
| DEFINITION | 602893J15F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5039019 5', | | | |
| ACCESSION | mRNA sequence. | | | |
| VERSION | BII07478 | | | |
| KEYWORDS | BII07478.1 GI:14558371 | | | |
| ORIGIN | EST. | | | |
| SOURCE | Mus musculus (house mouse) | | | |
| ORGANISM | Mus musculus | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| AUTHORS | Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus. | | | |
| TITLE | 1 (bases 1 to 838) | | | |
| JOURNAL | NIH-MGC http://mgc.nci.nih.gov/. | | | |
| COMMENT | National Institutes of Health, Mammalian Gene Collection (MGC) | | | |
| | Unpublished (1999) | | | |
| | Contact: Robert Strausberg, Ph.D. | | | |
| | Email: cgabbs-remail.nih.gov | | | |
| | Tissue Procurement: Gilbert Smith, Ph.D. | | | |
| | cDNA Library Preparation: Life Technologies, Inc. | | | |
| | DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) | | | |
| | Clone distribution: MGC clone distribution information can be | | | |
| | found through the I.M.A.G.E. Consortium/BLN at: | | | |
| | http://image.llnl.gov | | | |

| Plate: | LLAM1107 | row: | m | column: | 04 |
|---|--|---|-----|---------|----|
| High quality sequence strip: 830. | | | | | |
| Location/Qualifiers | | | | | |
| FEATURES | Source | 1. 838 | | | |
| | /organism="Mus musculus" | | | | |
| | /mol_type="mRNA" | | | | |
| | /strain="CZSCH 11" | | | | |
| | /db_xref="taxon:10090" | | | | |
| | /clone="IMAGE:5039019" | | | | |
| | /tissue_type="spontaneous tumor, metastatic to mammary. | | | | |
| | /stem_cell_origin="Stem cell origin." | | | | |
| | /lab_host="DH10B" | | | | |
| | /clone_1ib="NCI CGAP Lu29" | | | | |
| | /note="Organ: lung; Vector: pCMV-SPOrt6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIH" | | | | |
| ORIGIN | | | | | |
| Alignment Scores: | 2.35e-85 | Length: | 838 | | |
| Pred. No.: | 945.50 | Matches: | 174 | | |
| Score: | 82.774 | Conservative: | 23 | | |
| Percent Similarity: | 73.114 | Mismatches: | 36 | | |
| Best Local Similarity: | 46.094 | Indels: | 5 | | |
| Query Match: | 4 | Gaps: | 3 | | |
| DB: | | | | | |
| US-10-077-111-13 (1-384) x BI107478 (1-838) | | | | | |
| QY | 14 | ASPVALASbNCyCySaLaPheSerPheSerLeuLeuAlaThrCySerLeuAspLyThr | 33 | | |
| Db | 3 | GAGGTACGCTGCGTGGCCCTTCGGCGTCCCTCGGACCTGCTCTTGACAAAGCC | 62 | | |
| QY | 34 | ILeArgLeuTySerLeuAArgPheThrGluLeuPProHsSerProLeuLyPheHs | 53 | | |
| Db | 63 | ATCCGCTGACCTCCCTAAGTACCTTGCTGAACCGCGTACTCCCGCTAAGTTCAC | 122 | | |
| QY | 54 | ThrTyraLaAlaHisCyAcCyAcPheSerProSerGlyHsIleLeuAlaSerCySer | 73 | | |
| Db | 123 | ACCTATGCTGTCACCTGCTGCTGTTCTCACCTTGAGACAGCTTTTGCAATCGTCTCG | 187 | | |
| QY | 74 | ThrAspDgLYThrThrValLeuTrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGlu | 93 | | |
| Db | 183 | ACAAGACGGAGCCAGCGTGTGAGAGCTGCACACGAGACACACCTGACCGTGTGAG | 242 | | |
| QY | 94 | GlnProSerGlySerProValArgValCyGlnPheSerProAspSerThrCysLeuAla | 111 | | |
| Db | 243 | CACCGCGGTGGCGCCCTGTGGCGCTGTGCTTTTCCACAGACTCTGCTTACCTTGGC | 302 | | |
| QY | 114 | SerGlyAlaAlaAspDgLYThrValValLeuTrpAsnAlaGlnSerTyrrLyLeuTyArg | 133 | | |
| Db | 303 | TCAAGGGCTGCGCATGATATCCGTTGCTTGTGAAATGACACAGACTTACAACTTATAGG | 362 | | |
| QY | 134 | CyGgLYSerValLyAspDgLYSerLeuAlaAlaCySaLaPheSerProAsnGlySerPhe | 153 | | |
| Db | 363 | TGTGTGATGTCAAGATAGTCATTTGGTGGCTGTGCTTTTCCCGAGTGAAGGCTTC | 422 | | |
| QY | 154 | PheValThrGlySerSerCyGgLYAspLeuThrValTrpAspAspLyMetArgCysLeu | 173 | | |
| Db | 423 | TTTGTCACTGGCTCCCTCAGGCGGAGACTTGAACAGTGTGGATGACAGAAATAGGTGTCTA | 482 | | |
| QY | 174 | HisSerGlyLyAlaHisAspLeuGlyLeThrTyCySaAspPheSerSerGlnProVal | 199 | | |
| Db | 483 | CACAGGAGAAAGCGACGATCTCGGATACCTGTCTCAGCTTTTCTTCCACACGCTTC | 542 | | |
| QY | 194 | SerAspDgLYGlnGlnTyLeuGlnPhePheArgLeuAlaSerCyGgLYGlnAspCyGln | 213 | | |
| Db | 543 | TCTGGGGAGAA---GGCTCCAGTCTTACCAAGTTGGGTGATGTGGTCAAGACTGGAA | 592 | | |
| QY | 214 | ValLyIleThrPLeValSerPheThrHsIleLeuAlaArgGlyThrGlnHisGlnLeu | 233 | | |
| Db | 600 | ATCAAACTCTGGGCTGTATTAACCGTGTGCTTGGCTTTGAATTAACATAT----- | 653 | | |

Qy 234 LysGlnPheThrGluAspTyrSerGluGluValSerThrTyrPheCysAla 251
Db 654 ---AAAAGCACATTAAGTGGGACCTGGCTTCT---TGCCCTTGCT 701

RESULT 12

CK306752 839 bp mRNA linear EST 01-MAR-2004
LOCUS SB02035A1D03.f1 normalized Keck-Tagu Library SB02 Taeniopygia
DEFINITION guttata cDNA clone SB02035A1D03.f1 5, mRNA sequence.
ACCESSION CK306752
VERSION CK306752.1 GI:44816326
KEYWORDS EST.
SOURCE Taeniopygia guttata
ORGANISM Taeniopygia guttata
Taxonomy: Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Passeriformes; Estrildidae;
Estrildinae; Taeniopygia.
REFERENCE 1 (bases 1 to 839)
Clayton,D.F., Arnold,A.P., Ball,G.F., Brenowitz,B., George,J.M.,
Mello,C.V., Wade,J., Replogle,K., Lewin,H., Band,M., Hernandez,A.
and Liu,L.
The Songbird Neurogenomics Initiative: An Evolving Public Resource
for Study of Genes, Brain, and Behavior
Unpublished (2004)
CONTACT: David F. Clayton
University of Illinois
8107 CLSL, 601 S. Goodwin, Urbana, IL 61801, USA
TEL: 217 244 3668
FAX: 217 244 1648
Email: dclayton@uiuc.edu
Base Calling/Quality Scores: PHRED from Washington University
Genome Center.
Vector Trimming: Cross match from Washington University Genome
Center PHRAP suite. Low quality bases (Phred score < 20) were
trimmed from both ends of the sequence by an in-house script.
This sequence is vector free and at least 200 bp in length. Funded
by PHS grant # R01 NS045264, 'Songbird Neurogenomics Initiative.'
PCR primers
FORWARD: TAATACGACTCCTATGAGG(T7)
BACKWARD: ATTAACCTCCTCAAG(T3)
Insert Length: 839 Std Error: 0.00
Plate: SB02035A1 row: D column: 03
Seq primer: TAATACGACTCCTATGAGG (T7)
High quality sequence stop: 839.
location/qualifiers
1. 839
/organism="Taeniopygia guttata"
/mol_type="mRNA"
/db_xref="taxon:59729"
/clone="SB02035A1D03.f1"
/tissue_type="brain"
/dev_stage="late embryo, post-hatch days 1, 10, 20, 45,
and adult (pooled)"
/lab_host="DH10B"
/clone_lib="normalized Keck-Tagu Library SB02"
/note="Organ: brain; Vector: pBS II SK(+); Site 1:
EcoRI(5' side of insert); Site 2: NotI (3' side of
insert); The library was constructed and normalized as
described by Bonaldo, M.F., Lennon, G. and Soares, M.B.
(1996), Genome Research 6(9): 791-806. An identifying tag
was added at the 3' during cDNA synthesis:
insertAAAAAAAAAAAAAAAAATGCA."

ORIGIN

Alignment Scores:

Pred. NO.: 1.24e-81 Length: 839
Score: 907.00 Matches: 161
Percent Similarity: 84.00% Conservative: 28
Best Local Similarity: 71.56% Mismatches: 36
Query Match: 44.31% Indels: 0
Gaps: 0

US-10-077-111-13 (1-384) x CK306752 (1-839)

Qy 1 MetValIysLeuIleHisThrLeuAlaAspHisGlyAspAspValancCysAlaPhe 20
Db 90 ATGGCGACATTAAATTCACATTAGCAGATCATATGATGATTTATTAATGCGCTTC 149

Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg 40
Db 150 TCATCTTCGTCGCTGGCTGACATGTTCTCTTGACAAAAACAATTCGCGCTTATTTCTTGAGC 209

Qy 41 AspPheThrGluLeuPheHisSerProLeuLysPheHisThrTyrAlaHisCysCys 60
Db 210 AACTTCGCTAGCTCCGCTGACTCGCGTGAAGGTCACCGTACCGCGCTGACTGCTGC 269

Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeu 80
Db 270 TGCTTCGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 329

Qy 81 TyrAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 330 TGGGACACCCGCGATGCGCCGAGGCTGCGCTGCTGAGCAGCCCGGAGCGCGCTGCTG 389

Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAspGlyThr 120
Db 390 AGGCTCTGCGGCTTCTGCGCCGAGCGCGCTATCTTGGCAGAGGCGAGCGAGTGCAGC 449

Qy 121 ValValLeuThrAsnAlaGlnSerTyrLysLeuTyrArgCysGlySerValIysAspGly 140
Db 450 GTGCTTCTTGAGACGTCGACATCATGAACTTACCGATCTGGGAAAGTTAAAGTGT 509

Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
Db 510 TCTTGTAGTGCTTGTCATTTCTCCCAATGAAACTTCTTGTCACTGATCATCAAGT 569

Qy 161 GlyAspLeuThrValTyrPheAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
Db 570 GGTGATTTAAACAATTGGGATGATAAATAGATGCTGCTGATATATGAAAAAGCAGATG 629

Qy 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200
Db 630 CTTCGCGCTTACCTGCTGTGATTTCTTTCACATCCAGATATCTGATAGTAAATGATGC 689

Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValIysIleTyrIleValSer 220
Db 690 AATTAATCTTCAGATGCTTCTCTGTCGACAGATATATATCAAACTCGGCTATTG 749

Qy 221 PheThrHisIleLeu 225
Db 750 TTTCGACAGATTTCCTTA 764

RESULT 13

CK792250 808 bp mRNA linear EST 25-FEB-2004
LOCUS ACENECOURT18674252 NIH MGC 230 Mus musculus cDNA clone
DEFINITION IMAGE:30845589 5', mRNA sequence.
ACCESSION CK792250
VERSION CK792250.1 GI:42804246
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Taxonomy: Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 808)
NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Daniela S. Gerhard, Ph.D.
COMMENT Contact: Daniela S. Gerhard, Ph.D.
National Cancer Institute / NIH
Bldg. 31 Km10A07 Bethesda, MD 20892
Email: cgaaps-r@mail.nih.gov
Tissue Procurement: Shiko Kimura/Atsushi Yamada, (NCI, CCR)
cDNA Library Preparation: Express Genomics
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

Db 2 GCTCTCGAGACATATTTGGCATTCGTTCACACAGATGGTACCCTGCTTATGAAAT 61
Qy 83 ThrGluanglyGlnMetLeuAlaValMetGluGlnProSerGlySerProValArgVal 102
Db 62 ACTGAATAATGACAGATGCTGGCAGTGTATGAAACAGCTTACTGACGCCCTGTAGAGGT 121
Qy 103 CysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValVal 122
Db 122 TGCAGATTTCCTCCAGACTCCACGCTGTGGCATCAGGGGAGCTGATGAACTGTGGT 181
Qy 123 LeuThrPsnAlaGlnSerThrValLeuThrArgCysGlySerValValAspGlySerLeu 142
Db 182 TTGGGATGACACAGTATACAAATATATATAGTGTATGTTAAAGATGCTCTTG 241
Qy 143 AlaAlaCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCysGlyAsp 162
Db 242 GCGGCAATGTCATTTTCTCTTAATGAAAGCTTTCTTGTACCTGGCTCTCATGTGGTAT 301
Qy 163 LeuThrValTPAspAspPlyMetArgCysLeuHisSerGlyValAlaHisAspLeuGly 182
Db 302 TTACAGATGTGGATGATTAATATGAGGTCTGTCATAGTGAAGAACATGATCTTGA 361
Qy 183 IleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeuGlnPhe 202
Db 362 ATTACCTGCTGCGATTTTCTTCAACAGCCAGTTTCTGATGGAACAAGGCTCTCAGTTT 421
Qy 203 PheArgLeuAlaSerCysGlyGlnAspCysGlnValValIleTPriLeValSerPheThr 222
Db 422 TTTCGACTGGCATATGTCATGAGGATTTGCCAAGTCAAAATTTGATGTTCTTTTACC 481
Qy 223 HisIleLeu 225
Db 482 CATATCTTA 490

RESULT 15
BI759505 696 bp mRNA linear EST 25-SEP-2001
LOCUS 603046888F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5187309 5',
DEFINITION mRNA sequence.
ACCESSION BI759505
VERSION BI759505.1 GI:15751083
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 696)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rgabbe-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLU)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLU at:
http://image.llnl.gov
Plate: LHAM11467 row: O column: 22
High quality sequence stop: 696.
Location/Qualifiers
1..696
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5187309"
/lab_host="MDH10B"
/clone_lib="NIH_MGC_116"
/note="Organ: pooled colon, kidney, stomach; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 3 colons, age 26 yo male, 49 yo
female, 71 yo male colon; 46 yo male kidney, and pool of 2

stomachs, 62 yo male and 70 yo female. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.4 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
023. Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:

| | | | |
|------------------------|----------|---------------|-----|
| Pred. No.: | 5,26e-79 | Length: | 696 |
| Score: | 880.00 | Matches: | 161 |
| Percent Similarity: | 100.00% | Conservative: | 2 |
| Best local Similarity: | 98.77% | Mismatches: | 0 |
| Query Match: | 42.99% | Indels: | 0 |
| DB: | 4 | Gaps: | 0 |

US-10-077-111-13 (1-384) x BI759505 (1-696)

Qy 63 SerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrValLeuTPAsn 82
Db 2 GCTCTCGAGACATATTTGGCATTCGTTCACACAGATGGTACCCTGCTTATGAAAT 61
Qy 83 ThrGluanglyGlnMetLeuAlaValMetGluGlnProSerGlySerProValArgVal 102
Db 62 ACTGAATAATGACAGATGCTGGCAGTGTATGAAACAGCTTACTGACGCCCTGTAGAGGT 121
Qy 103 CysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThrValVal 122
Db 122 TGCAGATTTCCTCCAGACTCCACGCTGTGGCATCAGGGGAGCTGATGAACTGTGGT 181
Qy 123 LeuThrPsnAlaGlnSerThrValLeuThrArgCysGlySerValValAspGlySerLeu 142
Db 182 TTGGGATGACACAGTATACAAATATATAGTGTATGTTAAAGATGCTCTTG 241
Qy 143 AlaAlaCysAlaPheSerProAspGlySerPhePheValThrGlySerSerCysGlyAsp 162
Db 242 GCGGCAATGTCATTTTCTCTTAATGAAAGCTTTCTTGTACCTGGCTCTCATGTGGTAT 301
Qy 163 LeuThrValTPAspAspPlyMetArgCysLeuHisSerGlyValAlaHisAspLeuGly 182
Db 302 TTACAGATGTGGATGATTAATATGAGGTCTGTCATAGTGAAGAACATGATCTTGA 361
Qy 183 IleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeuGlnPhe 202
Db 362 ATTACCTGCTGCGATTTTCTTCAACAGCCAGTTTCTGATGGAACAAGGCTCTCAGTTT 421
Qy 203 PheArgLeuAlaSerCysGlyGlnAspCysGlnValValIleTPriLeValSerPheThr 222
Db 422 TTTCGACTGGCATATGTCATGAGGATTTGCCAAGTCAAAATTTGATGTTCTTTTACC 481
Qy 223 HisIleLeu 225
Db 482 CATATCTTA 490

Search completed: February 5, 2005, 14:52:37
Job time : 4214 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 05:07:50 ; Search time 246 Seconds
(without alignments)
8460.748 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272

Sequence: 1 gaatcgcttcaccctgcg.....acacaaagtaagaattc 1272

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database:

Issued Patents NA: *
1: /cgn2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgn2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgn2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgn2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgn2_6/ptodata/1/ina/PTUS COMB.seq: *
6: /cgn2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|---------|----|----------------------|
| 1 | 964.2 | 75.8 | 1844 | 4 | US-09-620-312D-768 |
| 2 | 49.6 | 3.9 | 7218 | 1 | US-08-232-463-14 |
| 3 | 45.8 | 3.6 | 7218 | 1 | US-08-232-463-14 |
| 4 | 44.8 | 3.5 | 1221 | 3 | US-08-965-600-2 |
| 5 | 44.8 | 3.5 | 1221 | 3 | US-09-489-506-2 |
| 6 | 43.8 | 3.4 | 1141 | 4 | US-09-806-708B-22 |
| 7 | 43.6 | 3.4 | 1141 | 4 | US-09-806-708B-22 |
| 8 | 43.6 | 3.4 | 937 | 4 | US-10-101-464A-251 |
| 9 | 40.4 | 3.2 | 162450 | 3 | US-09-345-882-1 |
| 10 | 40 | 3.1 | 2152 | 1 | US-08-188-582-17 |
| 11 | 40 | 3.1 | 2152 | 1 | US-08-646-715-17 |
| 12 | 40 | 3.1 | 246240 | 2 | US-08-724-394A-20 |
| 13 | 40 | 3.1 | 246240 | 2 | US-08-724-394A-20 |
| 14 | 40 | 3.1 | 246240 | 2 | US-08-724-394A-22 |
| 15 | 39.6 | 3.1 | 5152 | 3 | US-09-690-364-10 |
| 16 | 39.6 | 3.1 | 1664976 | 4 | US-08-916-421B-1 |
| 17 | 39.6 | 3.1 | 1664976 | 4 | US-09-692-570-1 |
| 18 | 38.2 | 3.0 | 2085 | 2 | US-08-283-917-8 |
| 19 | 38.2 | 3.0 | 2085 | 2 | US-08-961-716-8 |
| 20 | 38 | 3.0 | 2142 | 4 | US-09-614-221A-331 |
| 21 | 37.8 | 3.0 | 601 | 4 | US-09-949-016-138114 |
| 22 | 37.8 | 3.0 | 601 | 4 | US-09-949-016-138115 |
| 23 | 37.8 | 3.0 | 601 | 4 | US-09-949-016-138116 |
| 24 | 37.8 | 3.0 | 112705 | 4 | US-09-949-016-138116 |
| 25 | 37.6 | 3.0 | 1464 | 4 | US-09-248-796A-1456 |
| 26 | 37.4 | 2.9 | 29357 | 4 | US-09-949-016-16676 |
| 27 | 37 | 2.9 | 732 | 4 | US-09-248-796A-2341 |

| | | | | | | |
|----|------|-----|--------|---|----------------------|-------------------|
| 28 | 36.8 | 2.9 | 1611 | 3 | US-09-302-769-13 | Sequence 13, Appl |
| 29 | 36.6 | 2.9 | 3127 | 4 | US-09-710-279-3502 | Sequence 3502, Ap |
| 30 | 36.6 | 2.9 | 6968 | 4 | US-09-710-279-759 | Sequence 759, Ap |
| 31 | 36.6 | 2.9 | 7215 | 3 | US-09-134-001C-627 | Sequence 627, Ap |
| 32 | 36.4 | 2.9 | 3081 | 4 | US-09-949-016-2466 | Sequence 2466, Ap |
| 33 | 36.2 | 2.8 | 1566 | 4 | US-09-248-796A-9260 | Sequence 9260, Ap |
| 34 | 35.8 | 2.8 | 832 | 4 | US-09-621-976-2813 | Sequence 2813, Ap |
| 35 | 35.8 | 2.8 | 899 | 4 | US-09-370-767-11026 | Sequence 11026, A |
| 36 | 35.8 | 2.8 | 1743 | 4 | US-09-248-796A-11015 | Sequence 11015, A |
| 37 | 35.6 | 2.8 | 1182 | 4 | US-09-248-796A-4568 | Sequence 4568, Ap |
| 38 | 35.6 | 2.8 | 4892 | 4 | US-09-976-594-167 | Sequence 167, Ap |
| 39 | 35.6 | 2.8 | 5158 | 4 | US-09-023-655-1347 | Sequence 1347, Ap |
| 40 | 35.6 | 2.8 | 640681 | 4 | US-09-790-988-1 | Sequence 1, Appl |
| 41 | 35.6 | 2.8 | 640681 | 4 | US-09-790-988-1 | Sequence 1, Appl |
| 42 | 35.6 | 2.8 | 2417 | 4 | US-09-254-776B-3 | Sequence 3, Appl |
| 43 | 35.4 | 2.8 | 3747 | 3 | US-09-690-364-17 | Sequence 17, Appl |
| 44 | 35.4 | 2.8 | 7028 | 4 | US-09-949-016-4194 | Sequence 4194, Ap |
| 45 | 35.4 | 2.8 | 7028 | 4 | US-09-949-016-4194 | Sequence 4194, Ap |

ALIGNMENTS

RESULT 1
US-09-620-312D-768
; Sequence 768, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyang
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Dmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662e1 Nucleic Acids and
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_FL_gense Version 1.0
; SEQ ID NO 768
; LENGTH: 1844
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (164) ..(1594)
US-09-620-312D-768

Query Match 75.8%; Score 964.2; DB 4; Length 1844;
Best Local Similarity 81.8%; Pred. No. 1.9e-307;
Matches 1252; Conservative 0; Mismatches 3; Indels 276; Gaps 1;

QY 12 TCACCTGCGCGGACCTGACCCGCGCGCTTGAGCGGATCCCGCGG 71
DB 65 TCACCTGCGCGGACCTGACCCGCGCGCTTGAGCGGATCCCGCGG 124


```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/965,600
/ FILING DATE: Herewith
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER:
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J,
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0416 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-855-0555
/ TELEFAX: 650-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1221 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KIDNOT02
/ CLONE: 194046
US-08-965-600-2
```

Query Match 3.5%; Score 44.8; DB 3; Length 1221;

Best Local Similarity 49.3%; Pred. No. 0.0011; Mismatches 147; Indels 3; Gaps 1;

Matches 146; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

```
QY 328 CAACAGATGATGACCTGCTCTATGGAATCTGAAAATGACAGATGCTGGCAGTATG 387
DB 579 CCATAGATGGAATCATCATATTTTGTATGCACTGGAATACTTCGATACCTGG 638
QY 388 AACAGCTAGTGGCAGCCCTGTGAGGTTTCCAGTTTCCCACTCCAGCTGTTGG 447
DB 639 AAGGCC---ATGCCATGCCCATTCGCTCTTGAACCTTTCCCGGACTCCAGCTCCTTG 695
QY 448 CATCAGGGGACGCTATGGAACCTGTGTTTGTGGAATGACAGTATCAATATATATA 507
DB 696 TCACCTCTTCAGATGATGCTTACATCAAGTCTATGATGATACAACTGCCAATTTGGCTG 755
QY 508 GATGTGTAGTGTAAAGATGCTCCTTGGCGGCAATGCTATTTCTCTAATGAAAGCT 567
DB 756 GCACGCTGAGCGGCATGCTCCTGGGTGCTGAACGTGATTCGTGCTGATGACATC 815
QY 568 TCTTTGTCACTGCTCTCTCATGTGTGATTTAAACAGTGTGGATTAATAATGAG 623
DB 816 ACTTGTTCACAGTTCGTCTGACAAAGTGTAAAGTTGGGATGTTGGAACGAGG 871
```

RESULT 5

```
US-09-489-506-2
/ Sequence 2, Application US/09489506
/ Patent No. 6465619
/ GENERAL INFORMATION:
/ APPLICANT: Bandman, Olga
/ APPLICANT: Lal, Preeti
/ APPLICANT: Corley, Neil C.
/ TITLE OF INVENTION: NEW TRANSDUCIN BETA-1 SUBUNIT
/ NUMBER OF SEQUENCES: 3
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Incyte Pharmaceuticals, Inc.
/ STREET: 3174 Porter Drive
/ CITY: Palo Alto
/ STATE: CA
/ COUNTRY: USA
/ ZIP: 94304
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
```

```
/ SOFTWARE: FastSeq for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/489,506
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/965,600
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Billings, Lucy J,
/ REGISTRATION NUMBER: 36,749
/ REFERENCE/DOCKET NUMBER: PF-0416 US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 650-855-0555
/ TELEFAX: 650-845-4166
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1221 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ LIBRARY: KIDNOT02
/ CLONE: 194046
US-09-489-506-2
```

Query Match 3.5%; Score 44.8; DB 3; Length 1221;

Best Local Similarity 49.3%; Pred. No. 0.0011; Mismatches 147; Indels 3; Gaps 1;

Matches 146; Conservative 0; Mismatches 147; Indels 3; Gaps 1;

```
QY 328 CAACAGATGATGACCACTGCTCTATGGAATCTGAAAATGACAGATGCTGGCAGTATG 387
DB 579 CCATAGATGGAATCATCATATTTTGTATGCACTGGAATACTTCGATACCTGG 638
QY 388 AACAGCTAGTGGCAGCCCTGTGAGGTTTCCAGTTTCCCACTCCAGCTGTTGG 447
DB 639 AAGGCC---ATGCCATGCCCATTCGCTCTTGAACCTTTCCCGGACTCCAGCTCCTTG 695
QY 448 CATCAGGGGACGCTATGGAACCTGTGTTTGTGGAATGACAGTATCAATATATATA 507
DB 696 TCACCTCTTCAGATGATGCTTACATCAAGTCTATGATGATACAACTGCCAATTTGGCTG 755
QY 508 GATGTGTAGTGTAAAGATGCTCCTTGGCGGCAATGCTATTTCTCTAATGAAAGCT 567
DB 756 GCACGCTGAGCGGCATGCTCCTGGGTGCTGAACGTGATTCGTGCTGATGACATC 815
QY 568 TCTTTGTCACTGCTCTCTCATGTGTGATTTAAACAGTGTGGATTAATAATGAG 623
DB 816 ACTTGTTCACAGTTCGTCTGACAAAGTGTAAAGTTGGGATGTTGGAACGAGG 871
```

RESULT 6

```
US-09-806-708B-22
/ Sequence 22, Application US/09806708B
/ Patent No. 6784342
/ GENERAL INFORMATION:
/ APPLICANT: The University of British Columbia
/ TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
/ FILE REFERENCE: 4810-58741
/ CURRENT APPLICATION NUMBER: US/09/806,708B
/ CURRENT FILING DATE: 2001-04-03
/ PRIOR APPLICATION NUMBER: US 60/147,133
/ PRIOR FILING DATE: 1998-08-04
/ NUMBER OF SEQ ID NOS: 23
/ SOFTWARE: PatentIn version 3.0
/ SEQ ID NO 22
/ LENGTH: 1141
/ TYPE: DNA
/ ORGANISM: Artificial sequence
/ FEATURE:
/ NAME/KEY: promoter
/ LOCATION: (1)..(1141)
```

OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEI promoters
US-09-806-708B-22

Query Match 3.4%; Score 43.8; DB 4; Length 1141;
Best Local Similarity 9.6%; Pred. No. 0.0023;
Matches 67; Conservative 279; Mismatches 341; Indels 9; Gaps 1,

| | | |
|----|---|------|
| QY | CGGATGCGCATTTTCTCTTAATGGAACCTTTGTCACAGGCTCCCATGTGGTAT | 597 |
| Db | 53 MSKSRKTTAAAMCYKRRYYNNKSHWKGWTKKMYBEAANNSTSBTHARRKMDKTLAYBM | 112 |
| QY | 598 TTAACAGTGTGGATGATATAAATGAAGTGTCTGCATAGTGAAGAAAGACATGATCTTGGAA | 657 |
| Db | 113 TMTMKMGITGRRHHYKWRBAMBVDYDHHYVTAANNATTTGCMDDDDKRTTMMKKONNA | 172 |
| QY | 658 TTACCTGTCTGCGATTTTCTTCAACAGCAGTTTCTGATGAGAAACAGCTCTTCAGTTT | 717 |
| Db | 173 TGMDDDTKYHHMMNNNGCBTVMVVRKYKTRDMSBKRMVYGBMMKMGMSVDYTYVWVMD | 232 |
| QY | 718 TTGCATCTGCATCATGTGTGAGAGATTGCCAAGTCAAAATTTGGATTGTCTTTTACC | 777 |
| Db | 233 MCKKKYVRWVLT-----RGRRRNYYVAMBTAHRRYNNGWTBAMAYRRMTNNNNN | 283 |
| QY | 778 ATATCTTAGCAAGGCGACAGAAATCATGCTGAAGCAATTTAACCAATGTGTCAGAG | 837 |
| Db | 284 NAKAMCKRAKYGWNRBAYNSTCTTWSKTTKVRISCAANNCRADAKKHMMKMSAAM | 343 |
| QY | 838 AGCTCGTCTCAACATGCGTTGTGTGCACAGATTTAAAGATCTTGTGGTATTTTCAAGA | 897 |
| Db | 344 GVVYNNNNNNNNWYTKAARHBARWDMVMSAKKMHANAHHSRKKWTEYKRTKTVNNNG | 403 |
| QY | 898 TGAATTAACATTGATGAGAAAAGAACTGTGAAATCTTACAAAAGAAAGTGTGCTGATG | 957 |
| Db | 404 TTMKKRWAMWYKMDMBGTYYNNNNNGRTYYGWTKKKKWYTYKKYKANNCKPRAMDHK | 463 |
| QY | 958 TGAATTTGATCTCTAGGACTGCGTATTAAGTCTGAGGAAAATTGAAGCTCAGAGA | 1017 |
| Db | 464 TCTNNNTTMMCKYUWNNCYWKSMTNGSHBAAAVYVYWMWMMRYAHANNNNMDYWK | 523 |
| QY | 1018 CCAAGGTAAATCCCTTCTTCAAGAAATCTCGATAATTAATGCTCAATTAATAGAG | 1077 |
| Db | 524 ACTYTKYBVCGRMMNNYAATYTKSSMYTSTRYKPKTKNSMRKBDTSMGRANNYAAAB | 583 |
| QY | 1078 AACTTAGAAAGATCCGCTCATCGCATCAGATGGCTATTTCATATGAAGAAGACATG | 1137 |
| Db | 584 HYGKMYNTRMBWSHTHBEBRAGAAHYWMBMYBAKCHCMKAWYAKKYAGAGGSSNNNN | 643 |
| QY | 1138 AAAATTGGATCAGCAAAAAGAAAGCTACAGTCCCATGACAAATTTGTCTTCTTACG | 1197 |
| Db | 644 NNNNNNNNNNNATCARDDYAAASRYAAMAAKMYUYKEAANNAAYYTHANNWGCWNNATD | 703 |
| QY | 1198 CGGTACTTACACCAATAGACTCGAATAATGGCCA | 1233 |
| Db | 704 TTRTMMKXNNNNNAGTVMKNNNNNNAKNSAAAKNVAA | 739 |

```

RESULT 7
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; CURRENT FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: US 60/147,133
; PRIOR FILING DATE: 1999-08-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA

```

```

: ORGANISM: Artificial sequence
:
: FEATURE:
: NAME/KEY: promoter
: LOCATION: (1)..(1141)
: OTHER INFORMATION: consensus sequence of A.c., L.a., and B.n. FAH1 promoters
US-09-806-708B-22

```

| | | | | | | | |
|-----------------------|--------|--------------|---------|------------|------|--------|-------|
| Query March | 3.4%; | Score | 43.6; | DB | 4; | Length | 1141; |
| Best Local Similarity | 13.2%; | Pred. No. | 0.0027; | | | | |
| Matches | 77; | Conservative | 192; | Mismatches | 309; | Indels | 5; |
| | | | | Gaps | 1; | | |

| | | | |
|----|------|---|------|
| Qy | 690 | TCGATGAGAAACAGGCTCTTCACTTTTTCGATGCGATCATGTGACAGATTGCCAA | 749 |
| Db | 1100 | TKMSATGTAMWTHAKGATMCWTWYTGTRRRCRTYAMRTYTTSSNANWSCATKEM | 1044 |
| Qy | 750 | GTCAAAATTGGATTGTTCTTTTACCCATATCTTACCAAGCGCAGAACATCAAGT | 809 |
| Db | 1040 | WTMKWATDKXRTIAYAWCAMBRNNNNMCATNGVAKSCATNNAMMYATTBMAVAAKAR | 981 |
| Qy | 810 | AAGCAATTACCGAAGATTGGTCAGAGAGAGTGGCTCTCAACATGGCTTTTGACACAGAT | 869 |
| Db | 980 | WAGNNMMYGAAGNKGMAAATMBWADTAGKCCNNNNNNWTTDVRMAKAKXNN | 921 |
| Qy | 870 | TTAAAGATCGTGTGGATTTTGAAGTGAATAACTAGTGGAAAAAGACTGTGAT | 929 |
| Db | 920 | NNNAWYACTNRAATNNKMAITHMKWTHGAHSKRTHHTRTCRRTKNNNNNARTYW | 861 |
| Qy | 930 | CTTACAAAGAAAGTCTGGCTGATGATTTTGAAATCTCTAGACTGCTAGTAA | 989 |
| Db | 860 | YHHAABRRMAAWWTRTNNNNNNNNNACBNTRTWABKHSWCNNNNNNNNNNNTWC | 801 |
| Qy | 990 | GTCGTGAGAAAATTGAAAGCTCAGACCAAGGTTAAATCCCTTTCTTCAAGAAATCCT | 1045 |
| Db | 800 | HYTANABCBCCRANNNNAARMAETCNNYMAATVTHTWCCYKTMWATWMTDWTMBT | 741 |
| Qy | 1050 | GATGAATTTATGTCCGATTAACAGAGACTTATGAAGAATCCGGTCATGCGATCAAT | 1109 |
| Db | 740 | TTTNNMTTSTNNNNNNNNNNMACTNNNNNNNNNNKAYVHAATNMGCMW----NNTDART | 686 |
| Qy | 1110 | GGCTATTTCATATGAAAAAGAACATGAGAAATTGGATCAGCAAAAAAGAAACGTACAGT | 1165 |
| Db | 685 | NNTTVMRRMWTNFKRTWYSTTRHHYTGATNNNNNNNNNNNNNNNNNSCCTBTMMTW | 626 |
| Qy | 1170 | CCATGACAAATCTTGTCTTCTCTTCAGCGGTACTTACCCAAATAGGACTGTGAAATG | 1223 |
| Db | 625 | TMKDEGTVTRKVKWBDITCTYUDVMADSWWVYANWRCBDVYTRNNYTKCSYAHSYW | 566 |
| Qy | 1230 | GCCATCAATGATGGCTGGAGACACACACACAAAGTAAAGAAATTC | 1272 |
| Db | 565 | YWSNNNAWYRYSARWSSWABRTTNNNNNNWWSGVRHRWAGTM | 523 |

RESULT 8
 US-10-101-464A-251
 Sequence 251, Application US/10101464A
 Patent No. 6768041
 GENERAL INFORMATION:
 APPLICANT: Strabala, Timothy
 APPLICANT: Nieuwenhuizen, Nicolaas
 APPLICANT: Higgins, Colleen M.
 TITLE OF INVENTION: Compositions Isolated from Plant Cells
 TITLE OR INVENTION: and Their Use in the Modification of Plant Cell Signaling
 FILE REFERENCE: 11000.1020c2
 CURRENT APPLICATION NUMBER: US/10/101.464A
 CURRENT FILING DATE: 2002-03-18
 PRIOR APPLICATION NUMBER: 09/704,302
 PRIOR FILING DATE: 2000-11-01
 PRIOR APPLICATION NUMBER: 09/728,986
 PRIOR FILING DATE: 1999-01-12
 PRIOR APPLICATION NUMBER: 60/162,866
 PRIOR FILING DATE: 1999-11-01
 PRIOR APPLICATION NUMBER: PCT/US00/00724
 PRIOR FILING DATE: 2000-01-11

```

; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 251
; LENGTH: 937
; TYPE: DNA
; ORGANISM: Eucalyptus grandis
US-10-101-464A-251

Query Match      3.4%; Score 43; DB 4; Length 937;
Best Local Similarity 51.9%; Pred. No. 0.0037;
Matches 97; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 1060 TATGTCCATTAAGTAAAGTATGAAAGTCCGGTCATCGCATCGATGCTATTTCAT 1119
DB 732 TGTGCCCCATCTTCAGGAATATATGATGATCTTCAGATTGGCGCGATGTTTCATT 791
QY 1120 ATGAAGAAGAGCAATGGAATAATGATTCAGCAAAAAGCAATGATCCCATGACCA 1179
DB 792 ATGAAGAGAGAGTATGCGGGAATGTTGACACGACGATACATCGCCAATGACCA 851
QY 1180 ATCTGTCTTCTTCAGCGGTACTTACCAATAAGACTTGAAAATGCCATCAATA 1239
DB 852 ACTTGAAGCTTATGCTATTTGATCTCATCTCCCAACGCGCTTACGCTGCGCAATTCAAG 911
QY 1240 GATGGCT 1246
DB 912 AGTGGCT 918

RESULT 9
US-09-345-882-1/c
; Sequence 1, Application US/09345882
; Patent No. 6399373
; GENERAL INFORMATION:
; APPLICANT: Bougenelerec, Lydie
; TITLE OF INVENTION: A NOCLEIC ACID ENCODING A RETINOBLASTOMA BINDING PROTEIN (RBP-7)
; FILE REFERENCE: GENSET.031A
; CURRENT APPLICATION NUMBER: US/09/345,882
; CURRENT FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: US 60/091,315
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/111,909
; PRIOR FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 140
; SOFTWARE: Patent.pm
; SEQ ID NO 1
; LENGTH: 162450
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72794
; OTHER INFORMATION: 5-124-273 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88073
; OTHER INFORMATION: 5-127-261 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 90842
; OTHER INFORMATION: 99-1437-325 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 93714
; OTHER INFORMATION: 5-128-60 : polymorphic base deletion of GT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97122
; OTHER INFORMATION: 99-1442-224 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 97152

; OTHER INFORMATION: 5-129-144 : polymorphic base deletion of T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99098
; OTHER INFORMATION: 5-130-257 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 99117
; OTHER INFORMATION: 5-130-276 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 103805
; OTHER INFORMATION: 5-131-395 : polymorphic base A or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 106940
; OTHER INFORMATION: 5-133-375 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108106
; OTHER INFORMATION: 5-135-155 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108149
; OTHER INFORMATION: 5-135-198 : polymorphic base insertion of GTTT
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108308
; OTHER INFORMATION: 5-135-357 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 108471
; OTHER INFORMATION: 5-136-174 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 124134
; OTHER INFORMATION: 5-140-120 : polymorphic base C or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 134362
; OTHER INFORMATION: 5-140-348 : polymorphic base insertion of A
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146328
; OTHER INFORMATION: 5-143-84 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 146345
; OTHER INFORMATION: 5-143-101 : polymorphic base A or C
; FEATURE:
; NAME/KEY: allele
; LOCATION: 150329
; OTHER INFORMATION: 5-145-24 : polymorphic base A or G
; FEATURE:
; NAME/KEY: allele
; LOCATION: 160031
; OTHER INFORMATION: 5-148-352 : polymorphic base G or T
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID30
; FEATURE:
; NAME/KEY: allele
; LOCATION: 72771..72817
; OTHER INFORMATION: polymorphic fragment 5-124-273 SEQ ID51
; FEATURE:
; NAME/KEY: allele
; LOCATION: 88050..88096
; OTHER INFORMATION: polymorphic fragment 5-127-261 SEQ ID31
```

| | |
|--------------------|--|
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 88050..88096 |
| OTHER INFORMATION: | polymorphic fragment 5-127-261 SEQ ID52 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 90819..90865 |
| OTHER INFORMATION: | complement polymorphic fragment 99-1437-325 SEQ ID49 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 90819..90865 |
| OTHER INFORMATION: | complement polymorphic fragment 99-1437-325 SEQ ID70 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 93690..93736 |
| OTHER INFORMATION: | polymorphic fragment 5-128-60 SEQ ID32 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 93690..93736 |
| OTHER INFORMATION: | polymorphic fragment 5-128-60 SEQ ID53 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 97099..97145 |
| OTHER INFORMATION: | polymorphic fragment 99-1442-224 SEQ ID50 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 97099..97145 |
| OTHER INFORMATION: | polymorphic fragment 99-1442-224 SEQ ID71 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 97130..97177 |
| OTHER INFORMATION: | polymorphic fragment 5-129-144 SEQ ID33 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 97130..97177 |
| OTHER INFORMATION: | polymorphic fragment 5-129-144 SEQ ID54 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 99075..99121 |
| OTHER INFORMATION: | polymorphic fragment 5-130-257 SEQ ID34 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 99075..99121 |
| OTHER INFORMATION: | polymorphic fragment 5-130-257 SEQ ID55 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 99094..99140 |
| OTHER INFORMATION: | polymorphic fragment 5-130-276 SEQ ID35 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 99094..99140 |
| OTHER INFORMATION: | polymorphic fragment 5-131-395 SEQ ID57 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 106918..106966 |
| OTHER INFORMATION: | polymorphic fragment 5-133-375 SEQ ID37 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 106918..106966 |
| OTHER INFORMATION: | polymorphic fragment 5-133-375 SEQ ID58 |
| FEATURE: | NAME/KEY: allele |
| LOCATION: | 108084..108130 |
| OTHER INFORMATION: | polymorphic fragment 5-135-155 SEQ ID38 |
| FEATURE: | |

```

NAME/KEY: allele
LOCATION: 108084..108130
OTHER INFORMATION: polymorphic fragment 5-135-155 SEQ ID59
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID39
FEATURE:
NAME/KEY: allele
LOCATION: 108127..108177
OTHER INFORMATION: polymorphic fragment 5-135-198 SEQ ID60
FEATURE:

Query Match      3.2%; Score 40.4; DB 3; Length 162450;
Best Local Similarity 52.4%; Pred. No. 1;
Matches 89; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Oy 1021 AGGTTAAATCCCTTTCCTTGAGAAATCCCGATGAATTTTATGTATGCCAATAACATGAGAAC 1080
Db 73713 AGGAAATATTATTATTACTGTAAGTAGTAACAAGTAAATATCCTCAAGAGAAAAGGAAAAAATAA 73655
Oy 1081 TTATGAAGAATCCGGCATGCATGCATGAGTGGCTATTATATATGAAAAAGAAATGAGAA 1140
Db 73653 ACATATATCCAAAATCTAATAAGAAAAGATTAATCATCTCAAGAGAAAAGGAAAAAATAA 7359
Oy 1141 ATTGATACGACAAAAGAAAAGTACGAATGCCATGACAAATCTGTTCCT 1190
Db 73593 AAAGAGCCAGCAACTCAAAAACGAGAAAAACAAATGAATATGTTCCT 73544

RESULT 10
US-08-188-582-17
Sequence 17, Application US/08188582
Patent No. 5534410
GENERAL INFORMATION:
APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanase, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: DATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAPE AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/188,582
FILING DATE: 28-JAN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/RAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277229
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2152 base pairs

```

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2112
US-08-188-582-17

Query Match 3.1%; Score 40; DB 1; Length 2152;
Best Local Similarity 48.3%; Pred. No. 0.065;
Matches 112; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 407 TGTGAGGCTTTGGCAGTTTCCCGACGCTCCAGCTGTTTGGCATCAGGGGCGAGCTGATG 466
DB 1599 TGTGAATTGTACAGAGATTCATCCAAATTCTAATTATGTTGCTACGGGCTCTGCAGACAG 1658
QY 467 AACTGTGCTTTGTGGAATGCACAGTCATCAAAATTATATAGATGTGTAGTCTTAAAGA 526
DB 1659 AACTGTGCGGCTCTGGGACGCTCTGAATGTGAAGATCTTCACTGACACAA 1718
QY 527 TGGCTCTTGGGGCGCATGTGCATTTCTCTTAATGGAAGCTTCTGTCTACGTGCTCTTC 586
DB 1719 GGGACCAATTCATTCCTTGACATTTCTCCCAATGGAGATTCCTGCTACAGAGCAAC 1778
QY 587 ATGTGTGATTTTAACAGTGTGGATGATTAATGAGGTGTCTGCATAGTGA 638
DB 1779 AGATGGACAGAGTCTCTTGGGATATGTGACATGTGTTGATGTTGAGAA 1830

RESULT 11
US-08-646-715-17
Sequence 17, Application US/08646715
Patent No. 5637686

GENERAL INFORMATION:

APPLICANT: Tjian, Robert
APPLICANT: Comai, Lucio
APPLICANT: Dynlacht, Brian D.
APPLICANT: Hoey, Timothy
APPLICANT: Ruppert, Siegfried
APPLICANT: Tanese, Naoko
APPLICANT: Wang, Edith
APPLICANT: Weinzierl, Robert O.J.
TITLE OF INVENTION: TATA-BINDING PROTEIN ASSOCIATED FACTORS,
TITLE OF INVENTION: NUCLEIC ACIDS ENCODING TAFs AND METHODS OF USE
NUMBER OF SEQUENCES: 36
CORRESPONDENCE ADDRESS:
ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT
STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-4187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/646,715
FILING DATE: 09-MAY-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/188,582
FILING DATE: 28-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Osman, Richard A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: A-57650-2/AJT/BAO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2152 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2112
US-08-646-715-17

Query Match 3.1%; Score 40; DB 1; Length 2152;
Best Local Similarity 48.3%; Pred. No. 0.065;
Matches 112; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 407 TGTGAGGCTTTGGCAGTTTCCCGACGCTCCAGCTGTTTGGCATCAGGGGCGAGCTGATG 466
DB 1599 TGTGAATTGTACAGATTCATCCAAATTCTAATTATGTTGCTACGGGCTCTGCAGACAG 1658
QY 467 AACTGTGCTTTGTGGAATGCACAGTCATCAAAATTATATAGATGTGTAGTCTTAAAGA 526
DB 1659 AACTGTGCGGCTCTGGGACGCTCTGAATGTGAAGATCTTCACTGACACAA 1718
QY 527 TGGCTCTTGGGGCGCATGTGCATTTCTCTTAATGGAAGCTTCTGTCTACGTGCTCTTC 586
DB 1719 GGGACCAATTCATTCCTTGACATTTCTCCCAATGGAGATTCCTGCTACAGAGCAAC 1778
QY 587 ATGTGTGATTTTAACAGTGTGGATGATTAATGAGGTGTCTGCATAGTGA 638
DB 1779 AGATGGACAGAGTCTCTTGGGATATGTGACATGTGTTGATGTTGAGAA 1830

RESULT 12
US-08-724-394A-20
Sequence 20, Application US/08724394A
Patent No. 5872237

GENERAL INFORMATION:

APPLICANT: Feder, John N.
APPLICANT: Kronmal, Gregory S.
APPLICANT: Lauer, Peter M.
APPLICANT: Ruddy, David A.
APPLICANT: Thomas, Winston
APPLICANT: Tsuchihashi, Zenta
APPLICANT: Wolff, Roger K.
TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
TITLE OF INVENTION: Sequences and Antibodies thereo
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA

ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/724,394A
FILING DATE: 01-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Filtz, Renee A.
REGISTRATION NUMBER: 35,136
REFERENCE/DOCKET NUMBER: 017957-000100
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-576-0200
TELEFAX: 415-576-0300
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:


```
/
/   LENGTH: 246240 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: not relevant
/   TOPOLOGY: not relevant
/   MOLECULE TYPE: cDNA
/   FEATURE:
/     NAME/KEY: misc_feature
/     LOCATION: 1..246240
/     OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-20

Query Match
Best Local Similarity 55.9%; Score 40; DB 2; Length 246240;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 773 TACCATATCTTACGAGCGCCACAGAACATCTAGCTGAAGCAATTTACCGAAGTTGTC 832
DB 45039 TACTCATGCGCTGATGAGTATTACCAAACTACCCCTGGGCCAATTAACAAACAGCAGAGC 45098

QY 833 AGAGAGGTGCTCTCAACATGCGCTTTGTGCACAGATTTAAAGATCTTGTGTATTTT 892
DB 45099 AAATGTGTTGTGTGTATTAATCTTTCTACACAAATACAGAAAAAGTATTCATGTT 45158

QY 893 CAAGATGATTAACATT 908
DB 45159 CAGCATTAATAAATT 45174

RESULT 13
US-08-724-394A-21
; Sequence 21, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
;   APPLICANT: Feder, John N.
;   APPLICANT: Krommal, Gregory S.
;   APPLICANT: Laufer, Peter M.
;   APPLICANT: Ruddy, David A.
;   APPLICANT: Thomas, Winston
;   APPLICANT: Tsuchihashi, Zenta
;   APPLICANT: Wolff, Roger K.
;   TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
;   TITLE OF INVENTION: Sequences and Antibodies Thereo
;   NUMBER OF SEQUENCES: 31
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
;     STREET: Two Embarcadero Center, 8th Floor
;     CITY: San Francisco
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94111-3834
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/724,394A
;     FILING DATE: 01-OCT-1996
;     CLASSIFICATION: 536
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Flits, Renee A.
;       REGISTRATION NUMBER: 35,136
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE: 415-576-0200
;         TELEFAX: 415-576-0300
;   INFORMATION FOR SEQ ID NO: 21:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 246240 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: not relevant
;       TOPOLOGY: not relevant
;     MOLECULE TYPE: cDNA
```

```
/
/   FEATURE:
/     NAME/KEY: misc_feature
/     LOCATION: 1..246240
/     OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-21

Query Match
Best Local Similarity 55.9%; Score 40; DB 2; Length 246240;
Matches 76; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 773 TACCATATCTTACGAGCGCCACAGAACATCTAGCTGAAGCAATTTACCGAAGTTGTC 832
DB 45039 TACTCATGCGCTGATGAGTATTACCAAACTACCCCTGGGCCAATTAACAAACAGCAGAGC 45098

QY 833 AGAGAGGTGCTCTCAACATGCGCTTTGTGCACAGATTTAAAGATCTTGTGTATTTT 892
DB 45099 AAATGTGTTGTGTGTATTAATCTTTCTACACAAATACAGAAAAAGTATTCATGTT 45158

QY 893 CAAGATGATTAACATT 908
DB 45159 CAGCATTAATAAATT 45174

RESULT 14
US-08-724-394A-22
; Sequence 22, Application US/08724394A
; Patent No. 5872237
; GENERAL INFORMATION:
;   APPLICANT: Feder, John N.
;   APPLICANT: Krommal, Gregory S.
;   APPLICANT: Laufer, Peter M.
;   APPLICANT: Ruddy, David A.
;   APPLICANT: Thomas, Winston
;   APPLICANT: Tsuchihashi, Zenta
;   APPLICANT: Wolff, Roger K.
;   TITLE OF INVENTION: Megabase Transcript Map: No. 5872237e1
;   TITLE OF INVENTION: Sequences and Antibodies Thereo
;   NUMBER OF SEQUENCES: 31
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP
;     STREET: Two Embarcadero Center, 8th Floor
;     CITY: San Francisco
;     STATE: CA
;     COUNTRY: USA
;     ZIP: 94111-3834
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/724,394A
;     FILING DATE: 01-OCT-1996
;     CLASSIFICATION: 536
;     ATTORNEY/AGENT INFORMATION:
;       NAME: Flits, Renee A.
;       REGISTRATION NUMBER: 35,136
;       TELECOMMUNICATION INFORMATION:
;         TELEPHONE: 415-576-0200
;         TELEFAX: 415-576-0300
;   INFORMATION FOR SEQ ID NO: 22:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH: 246240 base pairs
;       TYPE: nucleic acid
;       STRANDEDNESS: not relevant
;       TOPOLOGY: not relevant
;     MOLECULE TYPE: cDNA
;     FEATURE:
;       NAME/KEY: misc_feature
;       LOCATION: 1..246240
;       OTHER INFORMATION: /note= "HLA-H.CONTIG"
US-08-724-394A-22
```

| | | | | |
|-----------------------|-------|----------------|--------|----------------|
| Query Match | 3.1% | Score 40; | DB 2; | Length 246240; |
| Best Local Similarity | 55.9% | Pred. No. 1.9; | | |
| Matches | 76; | Conservative | 0; | Mismatches 60; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

| | | | |
|----|-------|---|--------|
| Oy | 773 | TACCCATATCTTGGCAAGGGGCAAGAACATCAGCTGAAAGCAATTACGGAAGATTGTC | 832 |
| | | | |
| Db | 45039 | TACTCAGCTGTGATGAAGTATTTACAAACATCACCTCGGGCCATATAAACAAAGCAGAGAC | 450980 |
| Oy | 833 | AGAGAGAGTGCCTCAACATGAGCTTTGTGCACAAAGTTTAAAGATCTTGTTATTTT | 892 |
| | | | |
| Db | 45099 | AAATGTGTTTGTGTGTATTAACCTTCTACACAAAATACAGAAAAAGTGAATTCAGTT | 45158 |
| Oy | 893 | CAAGATGAATTAACATT | 908 |
| | | | |
| Db | 45159 | CAGCATAAATTAATTT | 45174 |

```

RESULT 15
US-09-690-364-10
: Sequence 10, Application US/09690364
: Patent No. 6468795
: GENERAL INFORMATION:
: APPLICANT: Hong Zhang
: APPLICANT: Andrew T. Watt
: TITLE OF INVENTION: ANTISENSE MODULATION OF APAF-1 EXPRESSION
: FILE REFERENCE: RTS-0190
: CURRENT APPLICATION NUMBER: US/09/690,364
: CURRENT FILING DATE: 2000-10-17
: NUMBER OF SEQ ID NOS: 100
: SEQ ID NO 10
: LENGTH: 5152
: TYPE: DNA
: ORGANISM: Mus musculus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (586)...(4302)
: US-09-690-364-10

```

| | | | | |
|-----------------------|--------|-----------------|-----------|----------------|
| Query Match | 3.1%; | Score 39.6; | DB 3; | Length 5152; |
| Best Local Similarity | 54.9%; | Pred. No. 0.16; | | |
| Matches | 78; | Conservative | 0; | Mismatches 64; |
| | | | Indels 0; | Gaps 0; |

| | | | |
|----|------|---|------|
| OY | 442 | GTTTGGCATCAGGGGGACGTCATGTAAGATGHTGGTTTGTGAATGCACAGTCATYACAAAT | 501 |
| Db | 3122 | GATTGCTTTCTTGTCATTTATGTAAGCGGTAAGGTGTGAATGTCTYACCGGAAGAA | 3781 |
| OY | 502 | TATATAGTGTGTAAGTGTAAAGTGGCTCCTTGCGCGGCATGTGCATTTTCTGCTATG | 561 |
| Db | 3182 | TAGAAAGACTTTACTTGTTCATCAGGGCAGATGCTTTTCTGTGCTATCTTCTTGATG | 3841 |
| OY | 562 | GAAGCTTCTTGTCACTGGCTC | 583 |
| Db | 3842 | CGACCAAGTTTCTCTTACCTTC | 3863 |

Search completed: February 5, 2005, 08:22:39
Job time : 252 secs

| Db | LOCUS | SEQUENCE | 180 bp | DNA | linear | PAT 21-JAN-2004 |
|------------|---|-------------|--------|-----|--------|-----------------|
| Db | TTGATGAGAAAGAACTGTTGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAAAATG 1 | | | | | |
| RESULT 2 | COI49443/ | COI49443 | 180 bp | DNA | linear | PAT 21-JAN-2004 |
| LOCUS | Sequence 19465 from Patent WO0157276. | | | | | |
| DEFINITION | COI49443 | | | | | |
| ACCESSION | COI49443.1 | GI:41156793 | | | | |
| VERSION | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | | | | | |
| ORGANISM | Homio sapiens (human) | | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R. | | | | | |
| TITLE | Human genome-derived single exon nucleic acid probes useful for | | | | | |
| JOURNAL | analysis of gene expression in human bone marrow | | | | | |
| | Patent: WO 0157276-A 19465 09-AUG-2001; | | | | | |
| FEATURES | Acemica, Inc. (US) | | | | | |
| source | Location/Qualifiers | | | | | |
| | 1..180 | | | | | |
| | /organism="Homio sapiens" | | | | | |
| | /mol type="unassigned DNA" | | | | | |
| | /db_xref="taxon:9606" | | | | | |
| | /note="MAP TO AC006501.5-EXPRESSED IN BONE MARROW, SIGNAL | | | | | |
| | = 0.81-EST HUMAN HIT: AL040518.1, EVALUE 1.00e-94-WT HIT: | | | | | |
| | U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE | | | | | |
| | 6.80e-02" | | | | | |
| ORIGIN | | | | | | |
| | Query Match 13.9%; Score 176.8; DB 6; Length 180; | | | | | |
| | Best Local Similarity 98.9%; Pred. N1.1e-36; | | | | | |
| | Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0; | | | | | |
| QY | 787 CAAGCGGACAGAAATCAGCTGGAAGCAATTACCGAAGATTGGCAGAGAGCTCT 846 | | | | | |
| Db | 180 CAAGCGGACAGAAATCAGCTGGAAGCAATTACCGAAGATTGGCAGAGAGCTCT 121 | | | | | |
| QY | 847 CAACATGCGCTTGTGCACAAAGATTTTAAAGATCTTGTGTAATTTTCAAGATGAATACA 906 | | | | | |
| Db | 120 CAACATGCGCTTGTGCACAAAGATTTTAAAGATCTTGTGTAATTTTCAAGATGAATACA 61 | | | | | |
| QY | 907 TTGATGAGAAAGAACTGTTGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAAAATG 966 | | | | | |
| Db | 60 TTGATGAGAAAGAACTGTTGAATCTTACAAAGAAAGCTGGCTGATGATTTGAAAAATG 1 | | | | | |
| RESULT 3 | COQ232723/ | COQ232723 | 180 bp | DNA | linear | PAT 21-JAN-2004 |
| LOCUS | Sequence 19562 from Patent WO0157273. | | | | | |
| DEFINITION | COQ232723 | | | | | |
| ACCESSION | COQ232723.1 | GI:41215941 | | | | |
| VERSION | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | | | | | |
| ORGANISM | Homio sapiens (human) | | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | |
| | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | | |
| REFERENCE | | | | | | |
| AUTHORS | 1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R. | | | | | |
| TITLE | HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR | | | | | |
| JOURNAL | ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO | | | | | |
| | 3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US | | | | | |
| | 60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03 | | | | | |
| | August 2000 (03.08.00)<150> GB 24263,6<151> 03 October 2000 | | | | | |
| | (03.10.00)<150> US 60/236,359<151> 27 September 2000 | | | | | |
| | (27.09.00)<150> US 60/234,687<151> 21 September 2000 | | | | | |
| | (21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170> | | | | | |
| | Molecular Dynamics Sequence Labeling Engine | | | | | |
| | Patent: WO 0157273-A 19562 09-AUG-2001; | | | | | |
| | Acemica, Inc. (US) | | | | | |
| | Location/Qualifiers | | | | | |
| | 1..180 | | | | | |
| | /organism="Homio sapiens" | | | | | |
| | | | | | | |

| | | | | |
|-----------------------|---|-------------|---|-----------------|
| | source | 1..180 | /organism="Homo sapiens" | |
| | | | /mol_type="unassigned DNA" | |
| | | | /db_xref="taxon:9606" | |
| | | | /note="MAP TO AC006501.5-EXPRESSED IN ADULT LIVER, SIGNAL = 0.72-EST HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT: U43339.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE 6.80e-02" | |
| ORIGIN | | | | |
| Query Match | 13.9%; Score 176.8; DB 6; | Length 180; | | |
| Best Local Similarity | 98.9%; Pred. No. 1.1e-36; | | | |
| Matches 178; | Conservative 0; Mismatches 2; Indels 0; Gaps 0; | | | |
| OY | 787 CAAGCGCACGAACATCAGCTGAGAACAATTTCACCAAGATTGGTCGAGAGAGGTCTCT | 846 | | |
| Dd | 180 CAAAGCCGCACGAACATCAGCTGAGAACAATTTCACCAAGATTGGTCGAGAGAGATGCT | 121 | | |
| OY | 847 CAACATGGCTTTGGCAACAAGATTAAAGATCTGTGGATTTTCAGATGAATAACA | 906 | | |
| Dd | 120 CAACATGGCTTTGGCAACAAGATTAAAGATCTGTGGATTTTCAGATGAATAACA | 61 | | |
| OY | 907 TTGATGAAAAAAGACTGTGTAATCTTCAAAAGAAGCTGGCTGATGATTTGAAAAATTG | 966 | | |
| Dd | 60 TTGATGAAAAAAGACTGTGTAATCTTCAAAAGAAGCTGGCTGATGATTTGAAAAATTG | 1 | | |
| RESULT 4 | | | | |
| CQ270754/c | 180 bp | DNA | linear | PAT 23-JAN-2004 |
| LOCUS | Sequence 19015 from Patent WO015727. | | | |
| DEFINITION | CQ270754 | | | |
| ACCESSION | CQ270754 | | | |
| VERSION | CQ270754.1 GI:41243358 | | | |
| KEYWORDS | . | | | |
| SOURCE | Homo sapiens (human) | | | |
| ORGANISM | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. Penn.S.G., Hanzel,D.K., Chen,W. and Rank,D.R. Human genome-derived single exon nucleic acid probes useful for analysis of gene expression in human fetal liver Parent: WO 015727-A 19015 03-AUG-2001; Aecomica, Inc. (US) Location/Qualifiers 1..180 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" /note="MAP TO AC006501.5-EXPRESSED IN FETAL LIVER, SIGNAL = 1.3-EST HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT: U43339.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE 6.80e-02" | | | |
| FEATURES | | | | |
| source | | | | |
| ORIGIN | | | | |
| Query Match | 13.9%; Score 176.8; DB 6; | Length 180; | | |
| Best Local Similarity | 98.9%; Pred. No. 1.1e-36; | | | |
| Matches 178; | Conservative 0; Mismatches 2; Indels 0; Gaps 0; | | | |
| OY | 787 CAAGCGCACGAACATCAGCTGAGAACAATTTCACCAAGATTGGTCGAGAGAGGTCTCT | 846 | | |
| Dd | 180 CAAAGCCGCACGAACATCAGCTGAGAACAATTTCACCAAGATTGGTCGAGAGAGATGCT | 121 | | |
| OY | 847 CAACATGGCTTTGGCAACAAGATTAAAGATCTGTGGATTTTCAGATGAATAACA | 906 | | |
| Dd | 120 CAACATGGCTTTGGCAACAAGATTAAAGATCTGTGGATTTTCAGATGAATAACA | 61 | | |
| OY | 907 TTGATGAAAAAAGACTGTGTAATCTTCAAAAGAAGCTGGCTGATGATTTGAAAAATTG | 966 | | |
| Dd | 60 TTGATGAAAAAAGACTGTGTAATCTTCAAAAGAAGCTGGCTGATGATTTGAAAAATTG | 1 | | |
| RESULT 5 | | | | |

CQ308036/c
LOCUS CQ308036 180 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 19141 from Patent WO0186003.
ACCESSION CQ308036
VERSION CQ308036.1 GI:41268613
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human lung
Patent: WO 0186003-A 19141 15-NOV-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. .180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN LUNG, SIGNAL =
0.88-BST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"
ORIGIN
Query Match 13.9%; Score 176.8; DB 6; Length 180;
Best Local Similarity 98.9%; Pred. No. 1.1e-36;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 787 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 846
180 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTGGTCAGAGAGATCTCT 121
Db 847 CAACATGCTTTGTGCAACAAGATTAAAGATCTTTGTGATTTTCAAGATGATACACA 906
120 CAACATGCTTTGTGCAACAAGATTAAAGATCTTTGTGATTTTCAAGATGATACACA 61
Qy 907 TTGATGAAAAAGAACTGTTGATTTTACAAAGAAAGTGGCTGATGATTTGAAAAATTG 966
60 TTGATGAAAAAGAACTGTTGATTTTACAAAGAAAGTGGCTGATGATTTGAAAAATTG 1
Db
RESULT 6
CQ344856/c
LOCUS CQ344856 180 bp DNA linear PAT 23-JAN-2004
DEFINITION Sequence 18950 from Patent WO0157275.
ACCESSION CQ344856
VERSION CQ344856.1 GI:41293927
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human brain
Patent: WO 0157275-A 18950 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. .180
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC006501.5-EXPRESSED IN BRAIN, SIGNAL =
0.73-BST_HUMAN HIT: AL040518.1, EVALUE 1.00e-94-NT HIT:
U43139.1, EVALUE 1.10e-01-SWISSPROT HIT: Q16760, EVALUE
6.80e-02"
ORIGIN
Query Match 13.9%; Score 176.8; DB 6; Length 180;

Best Local Similarity 98.9%; Pred. No. 1.1e-36;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 787 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 846
180 CAAGGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGATTGGTCAGAGAGATCTCT 121
Db 847 CAACATGCTTTGTGCAACAAGATTAAAGATCTTTGTGATTTTCAAGATGATACACA 906
120 CAACATGCTTTGTGCAACAAGATTAAAGATCTTTGTGATTTTCAAGATGATACACA 61
Qy 907 TTGATGAAAAAGAACTGTTGATTTTACAAAGAAAGTGGCTGATGATTTGAAAAATTG 966
60 TTGATGAAAAAGAACTGTTGATTTTACAAAGAAAGTGGCTGATGATTTGAAAAATTG 1
Db
RESULT 7
CQ058095/c
LOCUS CQ058095 132 bp DNA linear PAT 19-JAN-2004
DEFINITION Sequence 8915 from Patent WO0157270.
ACCESSION CQ058095
VERSION CQ058095.1 GI:41032601
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
Human genome-derived single exon nucleic acid probes useful for
analysis of gene expression in human breast and hbl 100 cells
Patent: WO 0157270-A 8915 09-AUG-2001;
Aeomica, Inc. (US)
FEATURES
Source
1. .132
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN HBL100, SIGNAL =
0.68-BST_HUMAN HIT: BB675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"
ORIGIN
Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 976 GACTGGGTAGTAAGCTGAGGAAATGAAAGAGCTCAGACCAAGTTAAATCCCTTT 1035
132 GACTGGGTAGTAAGCTGAGGAAATGAAAGAGCTCAGACCAAGTTAAATCCCTTT 73
Db 1036 CTTCAGGAATTCCTGATGATTTATATGTCATACTAGAGAACTTATGAAAGATCCGG 1095
72 CTTCAGGAATTCCTGATGATTTATATGTCATACTAGAGAACTTATGAAAGATCCGG 13
Qy 1096 TCATGCGATCAG 1107
12 TCATGCGATCAG 1
Db
RESULT 8
CQ077416/c
LOCUS CQ077416 132 bp DNA linear PAT 20-JAN-2004
DEFINITION Sequence 13216 from Patent WO0157278.
ACCESSION CQ077416
VERSION CQ077416.1 GI:41047285
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1

/note="MAP TO AC009307.1-EXPRESSED IN BT474, SIGNAL =
0.7-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN

Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 GACTGCGTAGTAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1

RESULT 12
CO30294/c 132 bp DNA linear PAT 21-JAN-2004
LOCUS Sequence 17133 from Patent WO0157273.
DEFINITION CO30294
ACCESSION CO30294
VERSION CO30294.1 GI:41213512
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN ADULT LIVER<130> PB 0004 WO

3<150> US 60/180,312<151> 04 February 2000 (04.02.00)<150> US
60/207,456<151> 26 May 2000 (26.05.00)<150> US 09/632,366<151> 03
August 2000 (03.08.00)<150> GB 24263.6<151> 03 October 2000
(03.10.00)<150> US 60/236,359<151> 27 September 2000
(27.09.00)<150> US 60/234,687<151> 21 September 2000
(21.09.00)<150> US 09/608,408<151> 30 June 2000 (30.06.00)<170>
Molecular Dynamics Sequence Listing Engine
Patent: WO 0157273-A 17133 09-AUG-2001;
Aeomica, Inc. (US)

JOURNAL

FEATURES

source 1. .132
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN ADULT LIVER, SIGNAL
= 0.73-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN

Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 GACTGCGTAGTAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1

RESULT 13

CO268431/c 132 bp DNA linear PAT 23-JAN-2004
LOCUS Sequence 16692 from Patent WO0157277.
DEFINITION CO268431
ACCESSION CO268431
VERSION CO268431.1 GI:41241035
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN FETAL LIVER
Patent: WO 0157277-A 16692 09-AUG-2001;
Aeomica, Inc. (US)

FEATURES

source 1. .132
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN FETAL LIVER, SIGNAL
= 0.78-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

ORIGIN

Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 976 GACTGCGTAGTAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGAAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGAGAAATTTATATGTCCAACTACAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1

RESULT 14

CO305465/c 132 bp DNA linear PAT 23-JAN-2004
LOCUS Sequence 16570 from Patent WO0166003.
DEFINITION CO305465
ACCESSION CO305465
VERSION CO305465.1 GI:41266042
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Penn, S.G., Hanzel, D.K., Chen, W. and Rank, D.R.
TITLE HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN LUNG
Patent: WO 0166003-A 16570 15-NOV-2001;
Aeomica, Inc. (US)

FEATURES

source 1. .132
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="MAP TO AC009307.1-EXPRESSED IN LUNG, SIGNAL =
0.76-EST HUMAN HIT: BE675766.1, EVALUE 2.00e-68-NT HIT:
X71133.1, EVALUE 1.20e+00-SWISSPROT HIT: P54860, EVALUE
1.00e+00"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 5, 2005, 05:27:56 ; Search time 757 Seconds
(without alignments)
9947.042 Million cell updates/sec

Title: US-10-077-111-12

Perfect score: 1272
Sequence: 1 gaattcgagcttcaccctgcg.....acaccaaagtaagaattc 1272

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 4887668

Minimum DB seq length: 0
Maximum DB seq length: 200

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: N_Geneseq_16Dec04:*
2: geneseqn1980s:*
3: geneseqn1990s:*
4: geneseqn2000s:*
5: geneseqn2001as:*
6: geneseqn2001bs:*
7: geneseqn2002as:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--|
| C 1 | 176.8 | 13.9 | 180 | 4 | ABA70710 Human foe |
| C 2 | 176.8 | 13.9 | 180 | 4 | AAI50879 Human foe |
| C 3 | 176.8 | 13.9 | 180 | 4 | AAI50879 Probe #19 |
| C 4 | 176.8 | 13.9 | 180 | 4 | AAI44908 Human bon |
| C 5 | 176.8 | 13.9 | 180 | 4 | AAI18959 Human bra |
| C 6 | 176.8 | 13.9 | 180 | 4 | ABSA4572 Human liv |
| C 7 | 132 | 10.4 | 132 | 4 | ABSA19150 Human gen |
| C 8 | 132 | 10.4 | 132 | 4 | ABSA12383 Probe #13 |
| C 9 | 132 | 10.4 | 132 | 4 | ABSA68387 Human foe |
| C 10 | 132 | 10.4 | 132 | 4 | AAI48602 Probe #17 |
| C 11 | 132 | 10.4 | 132 | 4 | ABSA50437 Human bre |
| C 12 | 132 | 10.4 | 132 | 4 | AAK42535 Human bon |
| C 13 | 132 | 10.4 | 132 | 5 | ABSA42143 Human liv |
| C 14 | 132 | 10.4 | 132 | 5 | AAI08924 Probe #89 |
| C 15 | 87 | 6.8 | 92 | 4 | ABSI6579 Human gen |
| C 16 | 41 | 3.2 | 41 | 4 | AAH36289 Human col |
| C 17 | 36 | 2.8 | 36 | 6 | AAI64919 Beta-tran |
| C 18 | 33 | 2.6 | 186 | 2 | AAI45080 Human RET |
| C 19 | 32.6 | 2.6 | 157 | 10 | AAI26096 Human ear- |
| C 20 | 29.6 | 2.3 | 136 | 4 | ABX85822 Corn ear- Aba70722 Human foe |

| | | | | | |
|----|------|-----|-----|---|--------------------|
| 21 | 29.6 | 2.3 | 136 | 4 | AAI50891 |
| 22 | 29.6 | 2.3 | 136 | 4 | AAK44920 |
| 23 | 29.6 | 2.3 | 136 | 4 | AAK18971 |
| 24 | 29.6 | 2.3 | 136 | 4 | ABSA4584 |
| 25 | 29.6 | 2.3 | 136 | 6 | ABSI9162 |
| 26 | 29.2 | 2.3 | 195 | 2 | AAH86637 |
| 27 | 29 | 2.3 | 180 | 4 | ABA89218 |
| 28 | 28.8 | 2.3 | 152 | 4 | AAI23721 |
| 29 | 28.8 | 2.3 | 152 | 4 | AAI49034 |
| 30 | 28.8 | 2.3 | 152 | 4 | ABSA0864 |
| 31 | 28.8 | 2.3 | 152 | 4 | ABSA35788 |
| 32 | 28.8 | 2.3 | 153 | 4 | AAI21599 |
| 33 | 28.8 | 2.3 | 153 | 4 | ABA66672 Human foe |
| 34 | 28.8 | 2.3 | 153 | 4 | AAI46886 |
| 35 | 28.8 | 2.3 | 153 | 4 | ABA48763 |
| 36 | 28.8 | 2.3 | 153 | 4 | ABA33735 |
| 37 | 28.8 | 2.3 | 153 | 4 | AAK40831 |
| 38 | 28.8 | 2.3 | 153 | 4 | AAK15102 |
| 39 | 28.8 | 2.3 | 153 | 4 | ABSA40410 |
| 40 | 28.8 | 2.3 | 153 | 5 | AAI07294 |
| 41 | 28.8 | 2.3 | 153 | 6 | ABSI4782 |
| 42 | 28.4 | 2.2 | 60 | 6 | ABN39870 |
| 43 | 28.2 | 2.2 | 194 | 3 | AAI27087 |
| 44 | 28 | 2.2 | 147 | 5 | ADL37494 |
| 45 | 28 | 2.2 | 147 | 5 | ADL72350 |

ALIGNMENTS

RESULT 1
ABA70710/c
ID ABA70710 standard; DNA, 180 BP.
XX
AC ABA70710;
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #19015.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe, ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WP1; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human foetal liver.
XX
PS Claim 4; SEQ ID NO 19015; 639bp + Sequence Listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human foetal liver. The

CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;
Best Local Similarity 98.9%; Pred. No. 2.4e-43;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACATCTGAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 846
DB 180 CAAGGCGCACAGAACATCTGAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 121

QY 847 CAACATGGCTTTGTGCGACAGATTTAAAGATCTTGGTATTTTCAAGATGATPACA 906
DB 120 CAACATGGCTTTGTGCGACAGATTTAAAGATCTTGGTATTTTCAAGATGATPACA 61

QY 907 TTGATGGAAAAAGAACTGTGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAAAATG 966
DB 60 TTGATGGAAAAAGAACTGTGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAAAATG 1

RESULT 2
AAI50879/c
ID AAI50879 standard; DNA; 180 BP.

XX AAI50879;

DT 17-OCT-2001 (first entry)

DE Probe #19565 used to measure gene expression in human placenta sample.

KM Probe: microarray; human; placenta; antenatal diagnosis;
KM genetic disorder; ss.

XX Homo sapiens.

PN W0200157272-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000663.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

XX Claim 25; SEQ ID NO 19565; 654bp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).
XX CC The present sequence is one such probe. The probes are useful for
XX CC producing a microarray for predicting, measuring and displaying gene
XX CC expression in samples derived from human placenta. The probes are useful
XX CC for antenatal diagnosis of human genetic disorders

XX Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;

Best Local Similarity 98.9%; Pred. No. 2.4e-43;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACATCTGAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 846
DB 180 CAAGGCGCACAGAACATCTGAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 121

QY 847 CAACATGGCTTTGTGCGACAGATTTAAAGATCTTGGTATTTTCAAGATGATPACA 906
DB 120 CAACATGGCTTTGTGCGACAGATTTAAAGATCTTGGTATTTTCAAGATGATPACA 61

QY 907 TTGATGGAAAAAGAACTGTGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAAAATG 966
DB 60 TTGATGGAAAAAGAACTGTGAATCTTACAAAGAAAGTGTGCTGATGATTTGAAAAATG 1

RESULT 3
AAK44908/c
ID AAK44908 standard; DNA; 180 BP.

XX AAK44908;

DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed single exon probe SEQ ID NO: 19465.

KM Human; bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukemia; lymphoma; myeloma; ss.

XX Homo sapiens.

PN W0200157276-A2.

PD 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000668.

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human bone marrow.

XX Example 4; SEQ ID NO 19465; 658bp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukemia and myeloma. The present sequence is one of
XX CC the probes of the invention

XX Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;
Best Local Similarity 98.9%; Pred. No. 2.4e-43;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACATCTGAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 846
DB 180 CAAGGCGCACAGAACATCTGAGCAATTTACCGAAGATTGGTCAGAGAGTGTCT 121

QY 847 CAACATGCTTTGTGCAACAAGATTAAAGATCTGTTGATTTTCAAGATGATAACA 906
DB 120 CAACATGCTTTGTGCAACAAGATTAAAGATCTGTTGATTTTCAAGATGATAACA 61
QY 907 TTGATGAAAAAGAACTGTTGAATCTTCAAAAAGAACTCGGCTGATGATTTGAAAATTG 966
DB 60 TTGATGAAAAAGAACTGTTGAATCTTCAAAAAGAACTCGGCTGATGATTTGAAAATTG 1

RESULT 4

AAK18959/c
ID AAK18959 standard; DNA; 180 BP.

AAK18959;

05-NOV-2001 (first entry)

Human brain expressed single exon probe SEQ ID NO: 18950.

Human; brain expressed exon; gene expression analysis; probe; microarray;
Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
68.

Homo sapiens.

MO200157275-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US000667.

04-FEB-2000; 2000US-0180312P.

26-MAY-2000; 2000US-0207456P.

30-JUN-2000; 2000US-00608408.

03-AUG-2000; 2000US-00632366.

21-SEP-2000; 2000US-0234687P.

27-SEP-2000; 2000US-0236359P.

04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-483446/52.

Single exon nucleic acid probes for analyzing gene expression in human
brains.

Example 4; SEQ ID NO 18950; 650pp + Sequence Listing; English.

The present invention provides a number of single exon nucleic acid
probes which are derived from genomic sequences expressed in the human
brain. They can be used to measure gene expression in brain cell samples,
which may enable the diagnosis and improved treatment of nervous system
diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
epilepsy and cancers. The present sequence is one of the probes of the
invention

Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;

Best Local Similarity 98.9%; Pred. No. 2.4e-43;

Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACTAGCTGAAGCAATTTACCGAAGATTGTCAGAGAGTGTCT 846

DB 180 CAAGGCGCACAGAACTAGCTGAAGCAATTTACCGAAGATTGTCAGAGAGATGTCT 121

QY 847 CAACATGCTTTGTGCAACAAGATTAAAGATCTTGTGATTTTCAAGATGATAACA 906

DB 120 CAACATGCTTTGTGCAACAAGATTAAAGATCTTGTGATTTTCAAGATGATAACA 61

QY 907 TTGATGAAAAAGAACTGTTGAATCTTCAAAAAGAACTCGGCTGATGATTTGAAAATTG 966

DB 60 TTGATGAAAAAGAACTGTTGAATCTTCAAAAAGAACTCGGCTGATGATTTGAAAATTG 1

RESULT 5
AB544572/c
ID AB544572 standard; DNA; 180 BP.

AB544572;

25-FEB-2003 (first entry)

Human liver single exon probe, SEQ ID NO 19562.

Human; single exon nucleic acid probe; liver; cirrhosis;
hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
coronary heart disease; 68.

Homo sapiens.

MO200157273-A2.

09-AUG-2001.

30-JAN-2001; 2001WO-US000664.

04-FEB-2000; 2000US-0180312P.

26-MAY-2000; 2000US-0207456P.

30-JUN-2000; 2000US-00608408.

03-AUG-2000; 2000US-00632366.

21-SEP-2000; 2000US-0234687P.

27-SEP-2000; 2000US-0236359P.

04-OCT-2000; 2000GB-00024263.

(MOLE-) MOLECULAR DYNAMICS INC.

Penn SG, Hanzel DK, Chen W, Rank DR;

WPI; 2001-488898/53.

Human genome-derived single exon nucleic acid probes useful for analyzing
gene expression in human adult liver.

Claim 4; SEQ ID NO 19562; 658pp; English.

The invention relates to a single exon nucleic acid probe (SENP) (I) for
measuring human gene expression in a sample derived from human adult
liver, comprising one of 13109 defined nucleotide sequences given in the
specification (or complements/ fragments). The probe hybridises at high
stringency to a nucleic acid molecule expressed in the human adult liver.
(I) may be used for predicting, measuring and displaying gene expression
in samples derived from human adult liver. The genes identified may be
involved in genetic liver diseases such as cirrhosis,
hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
associated with coronary heart disease. AB525011-AB551005 represent human
liver single exon nucleic acid probes of the invention. Note: The
sequence information for this patent does not appear in the printed
specification but was obtained in electronic format directly from WIPO at
ftp.wipo.int/pub/published_pct_sequences

Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;

Query Match 13.9%; Score 176.8; DB 4; Length 180;

Best Local Similarity 98.9%; Pred. No. 2.4e-43;

Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 787 CAAGGCGCACAGAACTAGCTGAAGCAATTTACCGAAGATTGTCAGAGAGTGTCT 846

DB 180 CAAGGCGCACAGAACTAGCTGAAGCAATTTACCGAAGATTGTCAGAGAGATGTCT 121

QY 847 CAACATGCTTTGTGCAACAAGATTAAAGATCTTGTGATTTTCAAGATGATAACA 906

DB 120 CAACATGCTTTGTGCAACAAGATTAAAGATCTTGTGATTTTCAAGATGATAACA 61

QY 907 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAAAGTCTGCGTGAATTTGAAAATTG 966
DB 60 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAAAGTCTGCGTGAATTTGAAAATTG 1

RESULT 6
ABS19150/c
ID ABS19150 standard; DNA; 180 BP.
XX
AC ABS19150;
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID NO 19141.
XX
KM Human; db; single exon probe; asthma; lung cancer; COPD; ILD;
KM Chronic obstructive pulmonary disease; interstitial lung disease;
KM Familial idiopathic pulmonary fibrosis; neurofibromatosis;
KM tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KM Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KM pulmonary histiocytosis; lymphangioleiomyomatosis; Karsagenet syndrome;
KM pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KM primary ciliary dyskinesia; pulmonary hypertension;
KM hyaline membrane disease; open reading frame; ORF.
XX
OS Homo sapiens.
XX
PN WO200186003-A2.
XX
PD 15-NOV-2001.
XX
PF 30-JAN-2001; 2001WO-US000665.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
PT Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
PS Claim 4; SEQ ID NO 19141; 634bp; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX; the novel set of probes which hybridise at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX collected from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of the
XX array; identifying exons in a eukaryotic genome, comprising (a)
XX algorithmically predicting at least one exon from genomic sequences of
XX the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsagenet syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WPI at ftp.wpi.int/pub/published_pcc_sequences
XX
SQ Sequence 180 BP; 49 A; 42 C; 25 G; 64 T; 0 U; 0 Other;
XX
Query Match 13.9%; Score 176.8; DB 6; Length 180;
Best Local Similarity 98.9%; Pred. No. 2,4e-43;
Matches 178; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
QY 787 CAAGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGTTGGTCAGAGAGTCGTCT 846
DB 180 CAAGCGCACAGAACATCAGCTGAAGCAATTTACCGAAGTTGGTCAGAGAGAGTCGTCT 121
XX
QY 847 CAACATGCGCTTTGTGCAACAAGTTTAAAGATCTTGTGATTTTCAAGATGAATTAAC 906
DB 120 CAACATGCGCTTTGTGCAACAAGTTTAAAGATCTTGTGATTTTCAAGATGAATTAAC 61
XX
QY 907 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAAAGTCTGCGTGAATTTGAAAATTG 966
DB 60 TTGATGAAAAGAACTGTTGAATCTTACAAAAGAAAGTCTGCGTGAATTTGAAAATTG 1

RESULT 7
AA13283/c
ID AA13283 standard; DNA; 132 BP.
XX
XX AA13283;
XX
AC AA13283;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #13216 for gene expression analysis in human cervical cell sample.
XX
KM Probe; human; microarray; gene expression; cervical epithelial cell;
KM cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 13216; 487bp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTCTGAGAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAAGTCTGAGAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATACAGACTTATGAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATACAGACTTATGAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 8
ID ABA68387/c
XX ABA68387 standard; DNA; 132 BP.
XX
AC ABA68387;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human foetal liver single exon nucleic acid probe #16692.
XX
KM Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000669.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-483447/52.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
PS Claim 4; SEQ ID NO 16692; 639bp + Sequence Listing; English.

XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a single exon nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTCTGAGAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGCGTAGTAAAGTCTGAGAAATTGAAAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATACAGACTTATGAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATACAGACTTATGAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 9
ID AA148602/c
XX AA148602 standard; DNA; 132 BP.
XX
AC AA148602;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #17288 used to measure gene expression in human placenta sample.
XX
KM Probe; microarray; human; placenta; antenatal diagnosis;
XX
OS genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488897/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
PS Claim 25; SEQ ID NO 17288; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENPs).
CC The present sequence is one such probe. The probes are useful for
CC producing a microarray for predicting, measuring and displaying gene
CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTGTGAGGAAATTGAGAGCTCAGACCAAGTTAAATCCCTT 1035
DB 132 GACTGCGTAGTAAAGTGTGAGGAAATTGAGAGCTCAGACCAAGTTAAATCCCTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTATATGTCCAATACTAGAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTATATGTCCAATACTAGAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 10
ABA50437/C
ID ABA50437 standard; DNA; 132 BP.
AC ABA50437;
DT 01-FEB-2002 (first entry)
XX
XX
DE Human breast cell single exon nucleic acid probe #9132.
XX
XX
KM Human; microarray; single exon probe; gene expression; breast; disease;
KM cancer; ss.
XX
OS Homo sapiens.
PN MO200157271-A2.
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US0000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
XX
PS Claim 4; SEQ ID NO 9132; 327bp + Sequence Listing; English.
XX
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring

CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC http://wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAAGTGTGAGGAAATTGAGAGCTCAGACCAAGTTAAATCCCTT 1035
DB 132 GACTGCGTAGTAAAGTGTGAGGAAATTGAGAGCTCAGACCAAGTTAAATCCCTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTATATGTCCAATACTAGAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTATATGTCCAATACTAGAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 11
AAK42535/C
ID AAK42535 standard; DNA; 132 BP.
AC AAK42535;
DT 06-NOV-2001 (first entry)
XX
XX
DE Human bone marrow expressed single exon probe SEQ ID NO: 17092.
XX
XX
KM Human; bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukemia; lymphoma; myeloma; ss.
XX
OS Homo sapiens.
PN MO200157276-A2.
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US0000668.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
XX
PS Example 4; SEQ ID NO 17092; 658bp + Sequence Listing; English.
XX
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is one of
CC the probes of the invention

XX SQ Sequence 132 BP, 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred.No.1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCAATTAAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCAATTAAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 12
ABS42143/C
ID ABS42143 standard; DNA; 132 BP.
XX AC ABS42143;
XX 25-FEB-2003 (first entry)
DT XX
DE Human liver single exon probe. SEQ ID NO 17133.
XX
KM Human; single exon nucleic acid probe; liver; cirrhosis;
KM hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
KM coronary heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
DR WPI; 2001-488898/53.
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human adult liver.
XX
PS Claim 4; SEQ ID NO 17133; 658bp; English.
XX
CC The invention relates to a single exon nucleic acid probe (SEN) (I) for
CC measuring human gene expression in a sample derived from human adult
CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (I) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic, liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed

CC CC Specification but was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 132 BP, 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 4; Length 132;
Best Local Similarity 100.0%; Pred.No.1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGCGTAGTAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTT 1035
DB 132 GACTGCGTAGTAAGTCTGAGGAAATTTGAAGAGCTCAGGACCAAGTTAAATCCCTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCAATTAAGAACTTATGAAAGATCCGG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCAATTAAGAACTTATGAAAGATCCGG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 13
AAI08924/C
ID AAI08924 standard; DNA; 132 BP.
XX AC AAI08924;
XX 09-OCT-2001 (first entry)
DT XX
DE Probe #8915 used to measure gene expression in human breast sample.
XX
KM Probe; human; breast disease; breast cancer; development disorder; ss;
KM inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX OS Homo sapiens.
XX PN WO200157270-A2.
XX PD 09-AUG-2001.
XX PF 29-JAN-2001; 2001WO-US000661.
XX PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-476286/51.
DR WPI; 2001-476286/51.
XX PT Novel single exon nucleic acid probe used to measuring gene expression in
PT a human breast.
XX
PS Claim 25; SEQ ID NO 8915; 322bp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridises at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and non-
CC carcinoma tumours. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
SQ
Query Match 10.4%; Score 132; DB 5; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGGTGTAAAGTCTGAGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGGTGTAAAGTCTGAGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATTAAGAACTTATGAAGATCCG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATTAAGAACTTATGAAGATCCG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 14
ABS16579/C
ID ABS16579 standard; DNA; 132 BP.
AC ABS16579;
XX
XX 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID NO 16570.
XX
XX Human; db; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
XX familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
XX Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX pulmonary histiocytosis; lymphangioleiomyomatosis; Karsagen syndrome;
XX pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX primary ciliary dyskinesia; pulmonary hypertension;
XX hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
XX
XX WO200186003-A2.
XX
XX 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US000665.
XX
XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX
XX Claim 4; SEQ ID NO 16570; 634p; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of probes
CC; the novel set of probes which hybridise at high stringency to a nucleic
CC acid expressed in the human lung; measuring gene expression in a sample
CC derived from human lung, comprising (a) contacting the array with a
CC collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of the
CC array; identifying exons in a eukaryotic genome, comprising (a)
CC algorithmically predicting at least one exon from genomic sequences of
CC the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene;
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene expression
CC analysis, and for identifying exons in a gene, particularly using human
CC lung derived mRNA and for the study of lung diseases such as asthma, lung
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
CC Karsagen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
CC present sequence is a single exon probe open reading frame of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 132 BP; 35 A; 29 C; 24 G; 44 T; 0 U; 0 Other;
Query Match 10.4%; Score 132; DB 6; Length 132;
Best Local Similarity 100.0%; Pred. No. 1.1e-29;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 976 GACTGGTGTAAAGTCTGAGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTT 1035
DB 132 GACTGGTGTAAAGTCTGAGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTT 73
QY 1036 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATTAAGAACTTATGAAGATCCG 1095
DB 72 CTTGAGGAATTCCTGATGAATTTATATGTCCTCAATTAAGAACTTATGAAGATCCG 13
QY 1096 TCATCGCATCAG 1107
DB 12 TCATCGCATCAG 1
RESULT 15
AAH36289
ID AAH36289 standard; cDNA; 92 BP.
AC AAH36289;
XX
XX 03-SEP-2001 (first entry)
XX
XX Human colon cancer antigen encoding cDNA SEQ ID NO:3371.
XX
XX Human colon cancer; colon cancer antigen; diagnosis; detection;
XX colorectal carcinoma; ss.
XX
XX Homo sapiens.
XX
XX WO200122920-A2.
XX
XX 05-APR-2001.
XX
XX 28-SEP-2000; 2000WO-US026524.
XX
XX

XX 29-SEP-1999; 99US-0157137P.
PR 03-NOV-1999; 99US-0163280P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Ruben SM, Barash SC, Birse CE, Rosen CA;
XX WPI; 2001-235357/24.
DR P-PSDB; AAG76884.
XX
PT Nucleic acid encoding 4277 human colon cancer-associated polypeptides,
XX useful for preventing, diagnosing and/or treating colorectal cancers.
XX
PS Claim 1; Page 5177-5178; 9803pp; English.
XX
CC AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patient's own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
SQ Sequence 92 BP; 27 A; 9 C; 23 G; 28 T; 0 U; 5 Other;

Query Match 6.8%; Score 87; DB 4; Length 92;
Best Local Similarity 94.6%; Pred. No. 5.3e-16;
Matches 87; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 591 GGTGATTAAACAGTGTGGAGTGAATAAAGGTGTCTGCACTAGTGAATAAAGCACATGAT 650
DB 1 GGTGATTAAACAGTGTGGAGTGAATAAAGGTGTCTGCACTAGTGAATAAAGCACATGAT 60
OY 651 CTTGGAAATTAACCTGCTGCGATTTTCTTCA 682
DB 61 CTTGGAAATTAACCTGCTGCGATTTTNTNANA 92

Search completed: February 5, 2005, 08:47:55
Job time : 760 secs

This Page Blank (uspto)


```
OY 72 CCCCCCTCTGAGGCTGTTTTCTTCAATAAAGAACTGGTGAATGATTCACACA 131
    |||||
Db 129 CCCCCCTCTGAGGCTGTTTTCTTCAATAAAGAACTGGTGAATGATTCACACA 188
    |||||
OY 132 TTAGCTGATCANTGATGACGATGTCACACTGCTGTCCTCTTCCCTTCCCTGGGCTACT 191
    |||||
Db 189 TTAGCTGATCANTGATGACGATGTCACACTGCTGTCCTCTTCCCTTCCCTGGGCTACT 248
    |||||
OY 192 TGTCTCTTGACAAAACAATTCCCTGTACTGCTTACGTAAGTGAAGTGCACAT 251
    |||||
Db 249 TGTCTCTTGACAAAACAATTCCCTGTACTGCTTACGTAAGTGAAGTGCACAT 308
    |||||
OY 252 TCTCCATTGAAGTTTCATACCTTATGCTGCTGCTGCTGCTGCTTCTCCCTCAGACAT 311
    |||||
Db 309 TCTCCATTGAAGTTTCATACCTTATGCTGCTGCTGCTGCTGCTTCTCCCTCAGACAT 368
    |||||
OY 312 ATTTTGGCATCGTGTTCAACAGATGTAACCACTGCTCTTATGAAATGTAATAATGACAG 371
    |||||
Db 369 ATTTTGGCATCGTGTTCAACAGATGTAACCACTGCTCTTATGAAATGTAATAATGACAG 428
    |||||
OY 372 ATGCTGGCAGTGTATGGAACAGCCTTATGAGGAGGCTGAGGGTTTGGCCAGTTTCCCA 431
    |||||
Db 429 ATGCTGGCAGTGTATGGAACAGCCTTATGAGGAGGCTGAGGGTTTGGCCAGTTTCCCA 488
    |||||
OY 432 GACTCCACGTTTGGCATCAGGGGAGCTGATGGAAGTGTGTTTGGAAATGACAG 491
    |||||
Db 489 GACTCCACGTTTGGCATCAGGGGAGCTGATGGAAGTGTGTTTGGAAATGACAG 548
    |||||
OY 492 TCTATCAAAATTATATAGATGTGTATGTTAAAGATGCTCTTGGGCGCATGTCATTT 551
    |||||
Db 549 TCTATCAAAATTATATAGATGTGTATGTTAAAGATGCTCTTGGGCGCATGTCATTT 608
    |||||
OY 552 TCTCCTAATGAAGCTCTTGTCACTGCTCCCTCATGAGTGAATTAACAGTGGGAGT 611
    |||||
Db 609 TCTCCTAATGAAGCTCTTGTCACTGCTCCCTCATGAGTGAATTAACAGTGGGAGT 668
    |||||
OY 612 GATTAATGAGTGTCTGATAGTGAAGAAAGACATGATCTTGAATTAACCTGCTGAT 671
    |||||
Db 669 GATTAATGAGTGTCTGATAGTGAAGAAAGACATGATCTTGAATTAACCTGCTGAT 728
    |||||
OY 672 TTTTCTTCAACAGCAGTTTCTGATGAGAAACAAGTCTTGAATTTTGCATGCGCATCA 731
    |||||
Db 729 TTTTCTTCAACAGCAGTTTCTGATGAGAAACAAGTCTTGAATTTTGCATGCGCATCA 788
    |||||
OY 732 TGTGTCAGAGTTGGCCAAAGTGAATTTGATGTTCTTTTCCCATATCTT----- 784
    |||||
Db 789 TGTGTCAGAGTTGGCCAAAGTGAATTTGATGTTCTTTTCCCATATCTTGAAGTTTTT 848
    |||||
OY 785 ----- 784
    |||||
Db 849 GAAATTAATAATAAAGTACACTGAGTGGGCACTGTCCTGTTCTGCTTGTGCTTTT 908
    |||||
OY 785 ----- 784
    |||||
Db 909 TCCCATGATGGGAGATGCTAGTCTCAGGGCTCAGTGAATAGTCTGTCATAGTATAGAT 968
    |||||
OY 785 ----- 784
    |||||
Db 969 ACTAATACTGAGATATATCTTCAACATTTGATCTCAGACACAGGATATGTCACAACCTGT 1028
    |||||
OY 785 ----- 784
    |||||
Db 1029 GCTTTTGACCTAATACCTTTTACTTGTCTACTGTTTCAATGGAACAAAGTGAACATC 1088
    |||||
OY 785 -----AGCAAGGCGACAGAAACATCAGCTGAAGCAA 815
    |||||
Db 1089 TGGCAATTGACCTGGAAACATTTTGGCAAGGCGCACAGAAACATCAGCTGAAGCAA 1148
    |||||
OY 816 TTATACGAAGATTGGTCAGAGAGTGTCTCAACATGGCTTTGTGTCACAAAGTTTAAAA 875
    |||||
Db 1149 TTATACGAAGATTGGTCAGAGAGTGTCTCAACATGGCTTTGTGTCACAAAGTTTAAAA 1208
    |||||
```

```
OY 876 GATCTTGTGGTATTTTCAAGATGAATTAACATTTAGTGAAGAAAGAACTGTTGAATCTTACA 935
    |||||
Db 1209 GATCTTGTGGTATTTTCAAGATGAATTAACATTTAGTGAAGAAAGAACTGTTGAATCTTACA 1268
    |||||
OY 936 AAAAGAAGCTGGCTGATGATTTGAATAATGGAATCTCTAGGACTGGCTGTAAAGGCTG 995
    |||||
Db 1269 AAAAGAAGCTGGCTGATGATTTGAATAATGGAATCTCTAGGACTGGCTGTAAAGGCTG 1328
    |||||
OY 996 AGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTCTTCAAGAAATTCCTGATGAA 1055
    |||||
Db 1329 AGGAAATTTGAAGAGCTCAGACCAAGTTAAATCCCTTCTTCAAGAAATTCCTGATGAA 1388
    |||||
OY 1056 TTTATATGTCATTAATCTAGAGAACTTATGAAGAATCCGTCATGCGATCAGATGCTAT 1115
    |||||
Db 1389 TTTATATGTCATTAATCTAGAGAACTTATGAAGAATCCGTCATGCGATCAGATGCTAT 1448
    |||||
OY 1116 TCATATGAAGAAAGCAATGCAAAATTTGATGACGAAAGAAAGAAAGCTCAAGTCCCATG 1175
    |||||
Db 1449 TCATATGAAGAAAGCAATGCAAAATTTGATGACGAAAGAAAGAAAGCTCAAGTCCCATG 1508
    |||||
OY 1176 ACAATCTTGTCTTCTTCTTCAAGGCTTACACCAATAGGACTGTGAAGATGCGCATC 1235
    |||||
Db 1509 ACAATCTTGTCTTCTTCTTCAAGGCTTACACCAATAGGACTGTGAAGATGCGCATC 1568
    |||||
OY 1236 AATGATGCTGAGACACACCAAAAGTAAA 1266
    |||||
Db 1569 AATGATGCTGAGACACACCAAAAGTAAA 1599
    |||||

RESULT 2
CR619104 1600 bp mRNA linear HTC 21-JUL-2004
LOCUS full-length cDNA clone CS01075YB08 of Placenta Cot 25-normalized
DEFINITION of Homo sapiens (human).
ACCESSION CR619104
VERSION CR619104.1 GI:50499911
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1600)
AUTHORS Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLES Full-length cDNA libraries and normalization.
JOURNAL Unpublished
REMARK Contact : Feng Liang Email : fliang@life.rockefeller.edu URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue
2 (bases 1 to 1600)

REFERENCE
Genoscope.
AUTHORS Direct Submission
TITLES Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
JOURNAL BP 191 91006 Evry cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source
1..1600
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS01075YB08"
/issue_type="Placenta Cot 25-normalized"
/plasmid="pCMVSPORT_6"

ORIGIN
Query Match 61.5%; Score 782; DB 3; Length 1600;
Best Local Similarity 77.3%; Pred. No. 1,6e-212;
Matches 1075; Conservative 0; Mismatches 180; Indels 135; Gaps 5;
```

```

Qy 12 TCACCTGCGGCGACGTGACCCGCGCCGCGGACCTTGAAGGCGATCCCGCGC 71
Db 13 TCACCTGCGGCGACGTGACCCGCGCCGCGGACCTTGAAGGCGATCCCGCGC 72
Qy 72 CCCCCGCTCGACAGGCTGTTTTCTTCAATAAAGAAACATGTAACCTGATTCACA 131
Db 73 CCCCCGCTCGACAGGCTGTTTTCTTCAATAAAGAAACATGTAACCTGATTCACA 132
Qy 132 TTAGCTATGATGAGTGAAGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 191
Db 133 TTAGCTATGATGAGTGAAGATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCT 192
Qy 192 TGCTCTTGGACAAACAAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 251
Db 193 TGCTCTTGGACAAACAAATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 252
Qy 252 TCTTCATTTGAAGTTTCATACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 311
Db 253 TCTTCATTTGAAGTTTCATACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
Qy 312 ATTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 371
Db 313 ATTTTGGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
Qy 372 ATGCTGGCATGATGTAACAGCCTAGTGGCAGCCCTGAGAGGTTTGCCAGTTTCCCA 431
Db 373 ATGCTGGCATGATGTAACAGCCTAGTGGCAGCCCTGAGAGGTTTGCCAGTTTCCCA 432
Qy 432 GACTCCACGCTGTTGGCATCAGGGGAGCTGATGGAATGTTGTTGTTGTTGTTGTTGTT 491
Db 433 GACTCCACGCTGTTGGCATCAGGGGAGCTGATGGAATGTTGTTGTTGTTGTTGTTGTT 492
Qy 492 TCATACAAATATATATAGTGTGTAAGTGTAAAGTGTCTCTTGGGCGCATGTCATTT 551
Db 493 TCATACAAATATATATAGTGTGTAAGTGTAAAGTGTCTCTTGGGCGCATGTCATTT 552
Qy 552 TCTCTTAATGGAAGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611
Db 553 TCTCTTAATGGAAGCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
Qy 612 GATTAATGAGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 671
Db 613 GATTAATGAGGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
Qy 672 TTTTCTTCAACGCGCATGTTCTGATGAGAAACAAGTCTTCAAGTTTTCAGCTGGCATCA 731
Db 673 TTTTCTTCAACGCGCATGTTCTGATGAGAAACAAGTCTTCAAGTTTTCAGCTGGCATCA 732
Qy 732 TGTGTGAGGATGTCAGGTCAGAAATTTGATGTTGTTTCTTACCATATCTTGAAGG 791
Db 733 TGTGTGAGGATGTCAGGTCAGAAATTTGATGTTGTTTCTTACCATATCTTGAAGG 792
Qy 792 CGCAGAGATCAGCTGAGC-----AAATT 818
Db 793 GAATTAATAATTAAGTACAGTGAAGGCACTGTGCTCTGTTTGGCTGTGCTTTT 852
Qy 819 ACCGAGATTGCTGAGAGAGCTC----- 842
Db 853 TCCCATGATGGCAGATGCTAGTCTCAGGGTCAGTGAATAGTGTATATATATAT 912
Qy 843 -----GTTCAACATG 854
Db 913 ACTAATACTGAAATATATCTTCAACATGTAAGTCAACAGAGATATATATATATAT 972
Qy 855 CTTTGTGACAAAG-----TTTAAAGATCTTGTGTATTTTCAAGATGAATATAT 908
Db 973 GCTTTTGACCTAATACCTTTTACTTGTCTAGTGTCTGATGACAAACAGTGAACATC 1032
Qy 909 GATGAGAAAGAACTGTTGAATCTTCAAAAGAAAGTCTGCTGATGATTTGAAATTTGAA 968
Db 1033 TGCGCAATTTGACCTGAGAAAGCTTTGCGCAAGGCGCAGACAGAAACATCACTGAAGCAA 1092
Qy 969 TCT--CTAGAGACTGCTAGTAAAGTCTGAGAGAAATTTGAAGAGCTCAGACCAAGGTTA 1026

```

```

Db 1093 TTTACCGAAGATTTGTCAGAGAGATGTCCTCAACATGCTTTGTGCAACAAGATTTTAAA 1152
Qy 1027 AATTCCTTTCTTCAAGAAATTCCTGATGATATTTATGTGTCATTAACATGAGACTTTGA 1086
Db 1153 GATCTTGTGTGTTATTTTCAAGATGAAATTAATGATGAGAAAGAACTGTTGAATCTTACA 1212
Qy 1087 AA-----GATCGGTCATGCTGATGAGTGGCTATTCATATGAAAGAAAGCAATG 1136
Db 1213 AAGGAAGTCTGGCTGATGATTTGAAATTTGATGCTGATTCATATGAAAGAAAGCAATG 1272
Qy 1137 GAAATTTGATGACGCAAAAGAAAGAAAGTCAAGTCCCAATGCAAAATCTTGTCTTCTTCA 1196
Db 1273 GAAATTTGATGACGCAAAAGAAAGAAAGTCAAGTCCCAATGCAAAATCTTGTCTTCTTCA 1332
Qy 1197 GCGGTACTTACCACTAATAGACTCTGAAATATGCGCATGATGATGCTGAGACAC 1256
Db 1333 GCGGTACTTACCACTAATAGACTCTGAAATATGCGCATGATGATGCTGAGACAC 1392
Qy 1257 CAAAAGTAAA 1266
Db 1393 CAAAAGTAAA 1402

```

```

RESULT 3
AL533510 941 bp mRNA linear EST 24-MAR-2004
LOCUS AL533510
DEFINITION CS0D0004YJ15 5-PRIME, mRNA sequence.
ACCESSION AL533510
VERSION AL533510.3 GI:45708442
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 941)
REFERENCE 1
AUTHORS Li W.B., Gruber C., Jesse J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 13, 2001 this sequence version replaced gi:31260591.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 1653.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?cs=CS0D0004YJ15&c=1653.r.
Location/Qualifiers
1. 941

```

```

FEATURES
source
1. 941

```

```

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0D0004YJ15"
/rna_type="ADULT BRAIN"
/dev_stage="adult"
/clone_lib="Homo sapiens ADULT BRAIN"
/notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pCMVSPORT 6
vector. Library was not normalized."

```

ORIGIN

```

Query Match 60.3%; Score 767.4; DB 1; Length 941;
Best Local Similarity 98.7%; Pred. No. 2e-208;
Matches 765; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

```


QY 12 TCACCTGCGGCGACGTGACCCGCCGCGGCGACCTTGAAGCGGATCCCGCG 71
Db 68 TCCCTGCGCGGCGACGTGACCCGCCGCGGCGACCTTGAAGCGGATCCCGCG 127
QY 72 CCCCCCTCTGAGGCGTCTTCTCAATTAAGAACATGGGAACGATTCACACA 131
Db 128 CCCCCCTCTGAGGCGTCTTCTCAATTAAGAACATGGGAACGATTCACACA 187
QY 132 TTAGCTGATCATGATGACGATGTCACACTGCTGTCCTCTCTCTCTCTCTCTCT 191
Db 188 TTAGCTGATCATGATGACGATGTCACACTGCTGTCCTCTCTCTCTCTCTCTCT 247
QY 192 TGTCTCTTGGACAAACAAATTCGCTGTACTCGTTACGATCTTACTGAATGCCACAT 251
Db 248 TGTCTCTTGGACAAACAAATTCGCTGTACTCGTTACGATCTTACTGAATGCCACAT 307
QY 252 TCTCCATTTGAAGTTTCACTTACTGCTGTCACACTGCTGTCCTCTCTCTCTCTCTCT 311
Db 308 TCTCCATTTGAAGTTTCACTTACTGCTGTCACACTGCTGTCCTCTCTCTCTCTCTCT 367
QY 312 ATTTTGGCATGCTGTCACAGATGATGACCACTGCTCTTACTGAATTAAGTGAACAG 371
Db 368 ATTTTGGCATGCTGTCACAGATGATGACCACTGCTCTTACTGAATTAAGTGAACAG 427
QY 372 ATGCTGCGACGATGATGACCACTGCTGTCACACTGCTGTCACACTGCTGTCACACT 431
Db 428 ATGCTGCGACGATGATGACCACTGCTGTCACACTGCTGTCACACTGCTGTCACACT 487
QY 432 GACTCCACGCTGTTGGATGAGGGGCGACGATGAGGAATGCTGTTTGTGGAATGACAG 491
Db 488 GACTCCACGCTGTTGGATGAGGGGCGACGATGAGGAATGCTGTTTGTGGAATGACAG 547
QY 492 TCATCAAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 551
Db 548 GCATCAAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607
QY 552 TCTCTCAATGAAAGCTTCTTGTCTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 611
Db 608 TCTCTCAATGAAAGCTTCTTGTCTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 667
QY 612 GATTAATGAGGATGCTGCTGATGAGGAAGCAATGATCTTGAATTAAGTGAATGAGGAT 671
Db 668 GATTAATGAGGATGCTGCTGATGAGGAAGCAATGATCTTGAATTAAGTGAATGAGGAT 727
QY 672 TTTTCTTCAAGCAGCTTCTGATGAGGAAGCAAGCTCTTCAAGTCTTCTTCAAGCAGCA 731
Db 728 TTTTCTTCAAGCAGCTTCTGATGAGGAAGCAAGCTCTTCAAGTCTTCTTCAAGCAGCA 787
QY 732 TGTGTCAGGATGTCACAAATTTGATGATGATGATGATGATGATGATGATGATGAT 786
Db 788 TGTGTCAGGATGTCACAAATTTGATGATGATGATGATGATGATGATGATGATGAT 842

RESULT 4
AL553333 889 bp mRNA linear EST 30-MAR-2004
LOCUS AL553333 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CSDD1075YB08 5-PRIME, mRNA sequence.
ACCESSION AL553333
VERSION AL553333.3 GI:45858102
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1. (bases 1 to 889)
AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31275147.
Contact: Genoscope
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 Evry cedex - FRANCE

Email: sequef@genoscope.cns.fr, Web: www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
1653.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?c=CSDD1075YB08A04Pic=1653.r.
FEATURES
source
1. 889
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="CSDD1075YB08"
/issue_type="PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 59.5%; Score 756.6; DB 1; Length 889;
Best Local Similarity 99.0%; Pred. No. 2,5e-205;
Matches 767; Conservative 4; Mismatches 3; Indels 1; Gaps 1;
QY 12 TCACCTGCGGCGACGTGACCCGCCGCGGCGACCTTGAAGCGGATCCCGCG 71
Db 13 TCASCTGSGCGGACGATGACCCGCCGCGGCGACCTTGAAGCGGATCCCGCG 72
QY 72 CCCCCCTCTGAGGCGTCTTCTCAATTAAGAACATGGGAACGATTCACACA 131
Db 73 CCCCCCTCTGAGGCGTCTTCTCAATTAAGAACATGGGAACGATTCACACA 132
QY 132 TTAGCTGATCATGATGACGATGTCACACTGCTGTCCTCTCTCTCTCTCTCTCT 191
Db 133 TTAGCTGATCATGATGACGATGTCACACTGCTGTCCTCTCTCTCTCTCTCTCT 192
QY 192 TGTCTCTTGGACAAACAAATTCGCTGTACTGCTGTCACACTGCTGTCACACTGCTGTCACACT 251
Db 193 TGTCTCTTGGACAAACAAATTCGCTGTACTGCTGTCACACTGCTGTCACACTGCTGTCACACT 252
QY 252 TCTCAATGAAAGCTTCTTGTCTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 311
Db 253 TCTCAATGAAAGCTTCTTGTCTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 312
QY 312 ATTTTGGCATGCTGTCACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 371
Db 313 ATTTTGGCATGCTGTCACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 372
QY 372 ATGCTGCGACGATGATGACCACTGCTGTCACACTGCTGTCACACTGCTGTCACACTGCTGTCACACT 430
Db 373 ATGCTGCGACGATGATGACCACTGCTGTCACACTGCTGTCACACTGCTGTCACACTGCTGTCACACT 432
QY 431 AGACTCCACGCTGTTGGATGAGGGGCGACGATGAGGAATGCTGTTTGTGGAATGACAG 490
Db 433 AGACTCCACGCTGTTGGATGAGGGGCGACGATGAGGAATGCTGTTTGTGGAATGACAG 492
QY 491 GTCATCAAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 550
Db 493 GTCATCAAAATTAATAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 552
QY 551 TTCTCCAAATGAAAGCTTCTTGTCTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 610
Db 553 TTCTCCAAATGAAAGCTTCTTGTCTCACTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 612
QY 611 TGATTAATGAGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 670
Db 613 TGATTAATGAGGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
QY 671 TTTTCTTCAAGCAGCTTCTGATGAGGAAGCAAGCTCTTCAAGTCTTCTTCAAGCAGCA 730

| | | | |
|-----------------------|---|--|---------------------------------------|
| Db | 673 | TTTTTCTTCAAGCAGCAGATTCTGATGTGAGGAACAAGCTTCTAGTTTTTTGCAGCTGCATC | 732 |
| Qy | 731 | ATGTGTCACAGATTGCCCAATCAAAATTGGATTGTTCTTTTACCATATCTTA | 785 |
| Db | 733 | ATGTGTCAGATTGGCCCAATCAAAATTGGATTGTTCTTTTACCATATCTTA | 787 |
| RESULT 5 | | | |
| BMS43484 | | | |
| LOCUS | | | |
| DEFINITION | BMS43484 | 1010 bp | mRNA |
| ACCESSION | AGENCOURT 6492519 | NIH_MGC_124 | Homo sapiens cDNA clone IMAGE:5726572 |
| VERSION | BMS43484 | | |
| KEYWORDS | BMS43484.1 | GI:18773915 | |
| SOURCE | EST. | | |
| ORGANISM | Homo sapiens (human) | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| AUTHORS | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| TITLE | 1 (bases 1 to 1010) | | |
| JOURNAL | NIH-MGC http://mgc.nci.nih.gov/ . | | |
| COMMENT | National Institutes of Health, Mammalian Gene Collection (MGC) | | |
| | Unpublished (1999) | | |
| | Contact: Robert Strausberg, Ph.D. | | |
| | Email: cgabs-remail.nih.gov | | |
| | Tissue Procurement: Invitrogen | | |
| | cDNA Library Preparation: Life Technologies, Inc. | | |
| | cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN) | | |
| | DNA Sequencing by: Agencourt Bioscience Corporation | | |
| | Clone distribution: MGC clone distribution information can be | | |
| | found through the I.M.A.G.E. Consortium/LNLN at: | | |
| | http://image.lnl.gov | | |
| | Plate: L1AM12719 | row: e | column: 05 |
| | High quality sequence start: 6 | | |
| | High quality sequence stop: 669. | | |
| FEATURES | | | |
| source | location/Qualifiers | | |
| | 1..1010 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="mRNA" | | |
| | /db_xref="taxon:9606" | | |
| | /clone="IMAGE:5726572" | | |
| | /issue_type="hippocampus" | | |
| | /lab_host="DH10B" | | |
| | /clone_idb="NIH_MGC_124" | | |
| | /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: EcoRV | | |
| | (destroyed); Site 2: NotI; RNA source male hippocampus, | | |
| | age 27. Library is oligo-dT primed and directionally | | |
| | cloned (EcoRV site is destroyed upon cloning). Average | | |
| | insert size 1.4 kb, insert size range 0.9-4 kb. Library is | | |
| | normalized and enriched for full-length clones and was | | |
| | constructed by C. Gruber (Invitrogen). Research Genetics | | |
| | tracking code 012." | | |
| ORIGIN | | | |
| Query Match | 56.3%; | Score 716.6; | DB 4; length 1010; |
| Best Local Similarity | 97.6%; | Pred. No. 7.9e-194; | |
| Matches 760; | Conservative 0; | Mismatches 14; | Indels 5; Gaps 3; |
| Qy | 12 | TCACCTCGGGGACACGTGACCCCGACCCCGCTGTGGACCTTGAAGGGGATCCCGCGG | 71 |
| Db | 87 | TCACCTCGGGGACACGTGACCCCGACCCCGCTGTGGACCTTGAAGGGGATCCCGCGG | 146 |
| Qy | 72 | CCCCCGCTCCGACAGGCTGTCTTTCTTCAATTAAGACATGTGGAATCTATTCACACA | 131 |
| Db | 147 | CCCCCGCTCCGACAGGCTGTCTTTCTTCAATTAAGACATGTGGAATCTATTCACACA | 206 |
| Qy | 132 | TTAGCTGATCATGTGAGCATGTCAACTGCTGTGCTTCTCTCTTTCCTCTTGACT | 191 |
| Db | 207 | TTAGCTGATCATGTGAGCATGTCAACTGCTGTGCTTCTCTCTTTCCTCTTGACT | 266 |
| Qy | 192 | TGCTTCCTTGGACAAACAAATTGCGCTGTACTGTGTAAGCTTACTGAACTGGCAAT | 251 |
| Db | 267 | TGCTTCCTTGGACAAACAAATTGCGCTGTACTGTGTAAGCTTACTGAACTGGCAAT | 326 |

| | | | | |
|----|--|-----|---|-----|
| Oy | | 252 | TTCTCATTGAAGTTCATTAACCTAATGCGTGCCAGCTGCCTGTTCCTCCCTTCAGGAAT | 311 |
| Db | | 327 | TCTCATATGAAGTTCATAACTAATGCTGCTCACTGCTGCTGTTCTCCCTTCAGGAAT | 386 |
| Oy | | 312 | ATTTTGGCATTCTGTTTCAACAGATGATGACACTGCTCTCATGGAAATGACAAAATGACAG | 371 |
| Db | | 387 | ATTTTGGCATTCTGTTTCAACAGATGATGACACTGCTCTCATGGAAATGACAAAATGACAG | 446 |
| Oy | | 372 | ATGCTGGCAGTATGAAACAGCTTAGTGGCACCCCTGTGAAGGTTTTGCCAGTTTCCCCA | 431 |
| Db | | 447 | ATGCTGGCAGTATGAAACAGCTTAGTGGCACCCCTGTGAAGGTTTTGCCAGTTTCCCCA | 506 |
| Oy | | 432 | GACCTCCAGTGTGGATACAGGGGCGACGTAGTGAACCTGTGGTTTTGGAAATGACAG | 491 |
| Db | | 507 | GACCTCCAGTGTGGATACAGGGGCGACGTAGTGAACCTGTGGTTTTGGAAATGACAG | 566 |
| Oy | | 492 | TCATACAAATTAATTAATGATGTGTAGTAAAGATGGCTCTTGGCGGCATGTGCATT | 551 |
| Db | | 567 | TCATACAAATTAATTAATGATGTGTAGTAAAGATGGCTCTTGGCGGCATGTGCATT | 626 |
| Oy | | 552 | TCTCTTAATGAGACCTCTTTTGTGACTGTGCTCTCATGTGTGATTTAACAGTGTGGAT | 611 |
| Db | | 627 | TCTCTTAATGAGACCTCTTTTGTGACTGTGCTCTCATGTGTGATTTAACAGTGTGGAT | 686 |
| Oy | | 612 | GATAAATAGAAGTCTGCAATGAGAAAAACA-CATGATTTGGAAATTAACCTGCTGGCA | 670 |
| Db | | 687 | GATAAATAGAAGTCTGCAATGAGAAAAACA-CATGATTTGGAAATTAACCTGCTGGCA | 746 |
| Oy | | 671 | TTTTTCTTCAACGACCACTTCTGATGAGAACAA--GATCTTCAAGTTTTTTCGACTGGCA | 728 |
| Db | | 747 | TTTTTCTTCCACGCGGTTCTGATGAGAACAAAGCTTCCAGTTTTTTCGAAATGGCA | 806 |
| Oy | | 729 | TCATGTGGTCAGGATTGGCAAGTAAAA--TTTGGATTTGTTCTTTAACCATATCTTA | 785 |
| Db | | 807 | TCATGTGGTCAGGATTGGCAAGTAAAAATTTGGAATTTGTTCTTTAACCATATCTTA | 865 |

RESULT_6
BI603184
DEFINITION
LOCUS
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

BI603184 858 bp mRNA linear EST 07-SEP-2001
60324982BP1 NIH_MGC_96 Homo sapiens CDNA clone IMAGE:5301559 5',
mRNA sequence.
BI603184
BI603184.1 GI:15496123
EST.
Homo sapiens (human)
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC Htbp://mgc.nci.nih.gov/
1 (bases 1 to 858)
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: rgs@nci-nim.nih.gov
Tissue Procurement: Miklos Palokovics, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHLRI), Shizaki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNLNL at:
http://image.lnl.gov
Plate: LHAM11763 row: h column: 08
High quality sequence stop: 756.
location/Qualifiers

1..858
/oranism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/cclone="IMAGE:5301559"
/tissue type="hypothalamus"

FEATURES
source

/lab_hosc="DH10B"
 /clone_1lb="NIH_MGC_96"
 /note="Organ: brain; Vector: pBluescript (modified
 pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
 (GTCGAG); Oligo-dT primed using primer
 5'-TTTTTTTTTTTTTTN-3', size-selected for average
 insert size 2.3 kb and normalized to R01 5. This is a
 primary library enriched for full-length clones and
 constructed using the Cap-trapper method (Garland, in
 preparation). Library constructed by M. Brownstein
 (NIMH/NHGR), National Institutes of Health. Note: this is
 a NIH_MGC Library."

ORIGIN

Query Match 52.2%; Score 664.2; DB 4; Length 858;
 Best Local Similarity 93.9%; Pred. No. 8.2e-179;
 Matches 758; Conservative 0; Mismatches 38; Indels 11; Gaps 6;

QY 12 TCACCTGCGGCGACGTGACCCGACCCCGTGGGCACTTGAGGCGGATCCCGCG 71
 |||||
 DB 52 TCACCTGCGGCGACGTGACCCGACCCCGTGGGCACTTGAGGCGGATCCCGCG 111
 |||||
 QY 72 CCCCCGCTCTGAGGCGTGTCTTCTCAATAAAGAACTGGAACGTGATCAGACA 131
 |||||
 DB 112 CCCCCGCTCTGAGGCGTGTCTTCTCAATAAAGAACTGGAACGTGATCAGACA 171
 |||||
 QY 132 TTAGCTGATCATGTGTGACGATGTCAACGTGTCCTCTCTCTTCCCTCTTGAGTACT 191
 |||||
 DB 172 TTAGCTGATCATGTGTGACGATGTCAACGTGTCCTCTCTCTTCCCTCTTGAGTACT 231
 |||||
 QY 192 TGTCTCTTGACAAACCAATTCGCTGACTCGTTACGTGACTTACTGAACTGCGACAT 251
 |||||
 DB 232 TGTCTCTTGACAAACCAATTCGCTGACTCGTTACGTGACTTACTGAACTGCGACAT 291
 |||||
 QY 252 TCTCCATTGAAGTTTCTTACTCTATGCTGTCACAGCTGCTTCTCTCCCTCCAGAGCAT 311
 |||||
 DB 292 TCTCCATTGAAGTTTCTTACTCTATGCTGTCACAGCTGCTTCTCTCCCTCCAGAGCAT 351
 |||||
 QY 312 ATTTTGGCATCGTGTCAACAGATGTGACACAGTGTCTTATGSAATATGAAATGACAG 371
 |||||
 DB 352 ATTTTGGCATCGTGTCAACAGATGTGACACAGTGTCTTATGSAATATGAAATGACAG 411
 |||||
 QY 372 ATGCTGGCACTGATGGAACAGCTGAGGAGCCCTGTGAGGGTTGCCAGTTTCCCA 431
 |||||
 DB 412 ATGCTGGCACTGATGGAACAGCTGAGGAGCCCTGTGAGGGTTGCCAGTTTCCCA 471
 |||||
 QY 432 GACTCCACGTTTGGCATAGGGGCGAGCTGATGGAACGTGTTTGTGGAATGACAG 491
 |||||
 DB 472 GACTCCACGTTTGGCATAGGGGCGAGCTGATGGAACGTGTTTGTGGAATGACAG 531
 |||||
 QY 492 TCATACCAATTTATATAGATGTGTAAGTAAAGATGGCTCTTGGCGGAGTGCATTT 551
 |||||
 DB 532 TCATACCAATTTATATAGATGTGTAAGTAAAGATGGCTCTTGGCGGAGTGCATTT 591
 |||||
 QY 552 TCTCTTAATGAAGCTTCTTGTCTACTGCTCTCATGTGTGATTTTAAACAGTGTGGAT 611
 |||||
 DB 592 TCTCTTAATGAAGCTTCTTGTCTACTGCTCTCATGTGTGATTTTAAACAGTGTGGAT 651
 |||||
 QY 612 GATTAATAATGA-GGTGTCTGATAGTGAATAAGCATGATCTTGGAATTAC---CTGCT 666
 |||||
 DB 652 GATTAATAATGAAGTGTCTGATAGTGAATAAGCATGATCTTGGAATTACCTGGCTCC 711
 |||||
 QY 667 GCGATTTTCTTACAGCAGCTTCTGA-TGGAGAAACAAGTCTTCA-GTTTTTGTGAC 723
 |||||
 DB 712 GATTTTCTTACAGCAGCTTCTGA-TGGAGAAACAAGTCTTCAAGTTTTTGTGAC 771
 |||||
 QY 724 TGGCATC-ATGTGTGATGATTTGCCAAGT-CAAAATTTGATTTTCTTTTAAACCATTA 780
 |||||
 DB 772 TGGCATCATGTGTGATGATTTGCCAAGTCAAAATTTGATTTTCTTTTAAACCATTA 831
 |||||
 QY 781 TCTTGAAGAGCGGACAGAACTATGAC 807
 |||||
 DB 832 TCTTGAAGAGCGGACAGAACTATGAC 858
 |||||

RESULT 7
 BI821143
 LOCUS
 DEFINITION
 60305050DF1 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5176086 5',
 mRNA Sequence.

ACCESSION
 BI821143
 VERSION
 BI821143.1 GI:15932693
 KEYWORDS
 EST.
 SOURCE
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
 1 (bases 1 to 792)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@mail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LLM11438 row: 1 column: 07
 High quality sequence stop: 751.

FEATURES

source

1..792
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5176086"
 /lab_hosc="DH10B"
 /clone_1lb="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 49.3%; Score 627.4; DB 4; Length 792;
 Best Local Similarity 96.4%; Pred. No. 2.9e-168;
 Matches 749; Conservative 0; Mismatches 16; Indels 12; Gaps 10;

QY 12 TCACCTGCGGCGACGTGACCCGACCCCGTGGGCACTTGAGGCGGATCCCGCG 71
 |||||
 DB 20 TCACCTGCGGCGACGTGACCCGACCCG-CCGCGGACCTTGAAAGCGGATCCCGCG 78
 |||||
 QY 72 CCCCCGCTCTGAGGCGTGTCTTCTCAATAAAGAACTGGAACGTGATCAGACA 131
 |||||
 DB 79 CCCCCGCTCTGAGGCGTGTCTTCTCAATAAAGAACTGGAACGTGATCAGACA 138
 |||||
 QY 132 TTAGCTGATCATGTGTGACGATGTCAACGTGTCCTCTCTTCCCTCTTGAGTACT 191
 |||||
 DB 139 TTAGCTGATCATGTGTGACGATGTCAACGTGTCCTCTCTTCCCTCTTGAGTACT 195
 |||||
 QY 192 TGTCTCTTGACAAACCAATTCGCTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTG 251
 |||||
 DB 196 CTGCTCTTGACAAACCAATTCGCTGTAAGTGTGTAAGTGTGTAAGTGTGTAAGTGTG 255
 |||||
 QY 252 TCTCCATTGAAGTTTCTTACTCTATGCTGTGCACTGCTGCTTCCCTCCAGAGCAT 311
 |||||
 DB 256 TCTCCATTGAAGTTTCTTACTCTATGCTGTGCACTGCTGCTTCCCTCCAGAGCAT 315
 |||||

| | | | |
|----|-----|--|-----|
| Oy | 312 | ATTTGGCATGCTGTTTCAACAGATGGTACCTGCTCCATGGAAATACGAAATGGACG | 371 |
| Db | 316 | A-TTTGGATGCTGTTTCAACAGATGGTACCTGCTCCATGGAAATACGAAATGGACG | 374 |
| Oy | 372 | ATGTGGCAGTGTATGAAACAGCTTAGTGAGGACCCCTGTGAGGGTGTGCCAGTTTCCCA | 431 |
| Db | 375 | ATGTGGCAGTGTATGAAACAGCTTAGTGAGGACCCCTGTGAGGGTGTGCCAGTTTCCCA | 434 |
| Oy | 432 | GACTCCACGTTGTTGGCATCAGGGGCACTGATGGACCTGTGGTTTGTGGAAATGCACG | 491 |
| Db | 435 | GACTCCACGTTGTTGGCATCAGGGGCACTGATGGAACTGTGGTTTGTGGAAATGCACG | 494 |
| Oy | 492 | TCATACAAATATATATAGATGTGTAGTGTAAAGATGGCTCCTTGGCGCATGTCAATT | 551 |
| Db | 495 | TCATACAAATATATATAGATGTGTAGTGTAAAGATGGCTCCTTGGCGCATGTCAATT | 554 |
| Oy | 552 | TCTCTAATGGAAGCTTCTTTGTCACTGGCTCCTC-ATGTGTATTTTAAACAGTGTGGGA | 610 |
| Db | 555 | TCTCTAATGGAAGCTTCTTTGTCACTGGCTCCTCAATGTGTATTTTAAACAGTGTGGGA | 614 |
| Oy | 611 | TGATTTAAATGAGTGTCTGCATATAGGAA-AAAGCATCATGATCTTGTGAAATTAACC-TGCTGC | 668 |
| Db | 615 | TGATTTAAATGAGTGTCTGCATATAGGAAATTAAGCTCATGATCTTGTGAAATTAACCCTTGCTGC | 674 |
| Oy | 669 | GATTTTTCCTTCAACGCCAGTTTC-TGATGAGACAAGGCTTTCAGTTTTTTCGAC-TGG | 726 |
| Db | 675 | GATTTTTCCTTCAACGCCAGTTTCCTTGATGAGAGACAAGGCTTTCAGTTTTTTCGACCTGG | 734 |
| Oy | 727 | CATCATGT-GGCAGAGTTGCACCAATGCA-ATTTGGATGTGTTTCTTTTAACCATAT | 781 |
| Db | 735 | CATCATGTGGGTCAAGATTGGCATGTCAACAATTTGAGCTGTACTTTTAACCATAT | 791 |

| FEATURES | SOURCE |
|------------|---|
| RESULT 8 | CN793033 |
| LOCUS | CN793033 |
| DEFINITION | 4127990 BARC 8BOV Bos taurus cDNA clone 8BOV_46H02 5', mRNA |
| ACCESSION | CN793033 |
| VERSION | CN793033.1 |
| KEYWORDS | GI:47689013 |
| SOURCE | EST. |
| ORGANISM | Bos taurus (cow) |
| | Bos taurus |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos. |
| | 1 (bases 1 to 728) |
| REFERENCE | Baumann,R.G., Baldwin,R.L., Sonstegard,T.S., Van Tassell,C.P. and Matukumalli,L.K. |
| AUTHORS | Construction and Analysis of a cDNA Library Generated From Intestinal Muscle and Epithelial Tissues of Holstein Cattle Unpublished (2004) |
| JOURNAL | Contact: Richard G. Baumann |
| COMMENT | Bovine Functional Genomics Lab ANRI |
| | BLDG 162: BARC-EAST, Beltsville, MD 20705, USA |
| | Tel: 3015048604 |
| | Fax: 3015048744 |
| | Email: rbaumann@nri.barc.usda.gov |
| | Single pass sequencing. Baases called and trimmed with phred 0.000925 using options -trim alt ' ' -trim fastaVector identified by cross match using options -mismatch 12 -minscore 12 |
| | Plate: 46 row: H column: 02 |
| | Seq primer: CCTATTGAGTGACACTGTATGAC |
| | High quality sequence stop: 728. |
| | Location/Qualifiers |
| | 1..728 |

```

FEATURES
  source
    1..728
    location/qualifiers
      /organism="Bos taurus"
      /mol_type="mRNA"
      /strain="Holstein"
      /db_xref="taxon:9913"
      /clone="8BDV_46H02"
      /sex="female"

```

```

/tissue_type="Epithelial, Muscle"
/dev_stage="Lactating, Neonatal"
/lab_host="DH10B_Tony"
/lab_idb="BARC_8BOY"
/clone_idb="Organ: Intestine; Vector: pCMVSPORT6.1; Site: 1;
NotI; Site: 2: SCOR1; Normalized cow CDNA intestinal
library in pCMVSPORT6.1, constructed from equimolar mRNA
pools derived from 5 sources, 4 lactating intestinal, 1
neonatal intestinal 4/5 Lactating, Proximal Duodenum,
Jejunum, Distal Ileum, Colon, 1/5 Neonatal, Proximal
Duodenum, Jejunum, Distal Ileum"

```

| | Query Match | 41.5%; | Score 527.8; | DB 7; | Length 728; |
|---------------------------------|--|--|---------------------|----------------|-----------------|
| | Best Local Similarity | 88.2%; | Pred. No. 1.1e-139; | | |
| | Matches 574; | Conservative | 0; | Mismatches 77; | Indels 0; |
| Qy | 87 | GCTGTTTTTCTTCAATTAAGAACATGGTGAACATGATTCACACATTAAGTGAATCATG | 146 | | |
| Db | 78 | GCAGGATATTTTCCAGTAAGAACATGGTGAACATTAATTCACATTAAGCAGATCATAT | 137 | | |
| Qy | 147 | GACATATGTCACCTGCTGCTTCTCTTTTCCCTCTTGGTACTTGCTCCTTGGACAA | 206 | | |
| Db | 138 | GATATATGTCACCTGCTGCTTCTCTTTTCCCTCTTGGTACTTGCTCCTTGGACAA | 197 | | |
| Qy | 207 | ACAATTGGCCCTGTAAGTGTGATGCTTGAACATGACACATTCACATTTGAAGTTT | 266 | | |
| Db | 198 | ACAATTGGCCCTGTAATTTCTGATGATCTTTACGAAATGACATATCTTCATTTGAAGTT | 257 | | |
| Qy | 267 | CATACCTATGCTGTCCACCTGCTGCTGTTTCTCCCTTCAGGACATATTTTGGCATGCT | 326 | | |
| Db | 258 | CACACCTATGCTGTCCACCTGCTGCTGTTTCTCCCTTCAGGACATATTTTGGCATGCT | 317 | | |
| Qy | 327 | TCAACAGATGATACCATCTGTCTTATGGAATCTGAATAATGACAGATGCTGGCATGAT | 386 | | |
| Db | 318 | TCAACAGATGATACCATCTGTCTTATGGAATCTGAATAATGACAGAGATTTTGGCAGTGA | 377 | | |
| Qy | 387 | GACAGCCCTAGTGGCAGCCCTGTAGGGTTTGGCAGATTTTCCCGACATCCAGCTGTTTG | 446 | | |
| Db | 378 | GACAGCCCTAGTGGTACCCCTGTAGAGATTTGGCGATTTTCCCGACATCCAGCTGTTTG | 437 | | |
| Qy | 447 | GCATCAGGGCAGCTGATGGAATCTGGGTTTGTGGAATGACAGTCATCAAAATTATAT | 506 | | |
| Db | 438 | GTGTCAAGGGCAGCTGATGGAATCTGGGTTTGTGGAATGACAGTCATCAAAATTATAT | 497 | | |
| Qy | 507 | AGATGTGTATGTGTAAGAAGTGCCTCTGGCGGCAATGCAATTTCTCCAAATGGAAGC | 566 | | |
| Db | 498 | AGATGTGTATGTGTAAGAAGTGCCTCTCTAGTGCCTGTGGGTTTCTCTCGGAAC | 557 | | |
| Qy | 567 | TTCTTTTGTCACTGCTCCTCATGTGTGATTTAACAAGTGGAGATGAATAAATGAGGTG | 626 | | |
| Db | 558 | CTCTTTGTCACTGCTCCTCATGTGTGATTTAACAAGTGGAGATGAATAAATGAGGTG | 617 | | |
| Qy | 627 | CTGCATAGTGAATAAAGACATGATCTTGAATTAACCTGTGCGATTTTCTTCAACAGCA | 686 | | |
| Db | 618 | CTGCACAGTGAATAAAGACATGATCTTGAATTAACCTGTGCGATTTTCTTCAACACCA | 677 | | |
| Qy | 687 | GTTCCTGATGGAACAAAGTCTTCAGTTTCTTGGACTGGAATCATGTGCT | 737 | | |
| Db | 678 | GTTCCTGATGGAACAAAGGCTTCAGTTTCTTGAATGATGGCATCATGTGCT | 728 | | |
| RESULT 9 | | | | | |
| LOCUS | A1189142 | 505 bp | mRNA | linear | EST 13-OCT-1998 |
| DEFINITION | qqd4a04.x1 Soares placent85c0yewe8x2bhp8t05w Homo sapiens cDNA | | | | |
| DESCRIPTION | clone IMAGE:1722702 3' similar to SM:T2D3 DROME P49846 | | | | |
| TRANSCRIPTION INITIATION FACTOR | TFIID 85 KD SUBUNIT 1, mRNA | | | | |
| SEQUENCE | A1189142 | | | | |
| VERSION | A1189142.1 GI:3740351 | | | | |
| KEYWORDS | EST. | | | | |
| SOURCE | Homo sapiens (human) | | | | |

QY 476 TTGTGGAATGCACAGTCATACAAATTAATATAGATGTGTAGTGTAAAGTGGCTCCTT 535
 181 TTGTGGAATGCACAGTCATACAAATTAATATAGATGTGTAGTGTAAAGTGGCTCCTT 240
 Db 536 GGGGGCATGTGCATTTTCTCTTAATGGAAGCTTTTGTGTCACGTGGCTCTCATGTGGTGA 595
 241 GGGGGCATGTGCATTTTCTCTTAATGGAAGCTTTTGTGTCACGTGGCTCTCATGTGGTGA 300
 QY 596 TTAAACAGTGTGGGATGATAAATAGAGTGTCTGCATATGTAAGAAAGACATGATCTTGG 655
 301 TTAAACAGTGTGGGATGATAAATAGAGTGTCTGCATATGTAAGAAAGACATGATCTTGG 360
 Db 361 AATTACCTGCTGCAGATTTTCTTCTTCAAGCCAGTTTCTGATGAGAACAGAGTCTTCAGTT 420
 QY 656 AATTACCTGCTGCAGATTTTCTTCTTCAAGCCAGTTTCTGATGAGAACAGAGTCTTCAGTT 715
 361 AATTACCTGCTGCAGATTTTCTTCTTCAAGCCAGTTTCTGATGAGAACAGAGTCTTCAGTT 420
 Db 716 TTTTGCACGTGCATCATGTGTGTGTCAGATTTGCCAAGTCAAAATTTGATTTTCTTTTAC 775
 421 TTTTGCACGTGCATCATGTGTGTGTCAGATTTGCCAAGTCAAAATTTGATTTTCTTTTAC 480
 QY 776 CCATATCTTAG 786
 Db 481 CCATATCTTAG 491

RESULT 11
 B1759505 696 bp mRNA linear EST 25-SEP-2001
 LOCUS 603046888F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5187309 5',
 DEFINITION mRNA sequence.
 ACCESSION B1759505
 VERSION B1759505.1 GI:15751083
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 696)
 NIH-MGC http://mgs.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 CONTACT: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.lnl.gov
 plate: LLM11467 row: 0 column: 22
 High quality sequence stop: 696.
 Location/Qualifiers
 1. 696
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5187309"
 /lab_host="DH10B"
 /clone_id="NIH_MGC_116"
 /note="Organ: pooled colon, kidney, stomach; Vector:
 PCWV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 3 colons, age 36 yo male, 49 yo
 female, 71 yo male colon; 46 yo male kidney, and pool of 2
 stomachs, 62 yo male and 70 yo female. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.4 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 023. Note: this is a NIH_MGC library."

FEATURES
 source

Query Match 38.2%; Score 486.2; DB 4; Length 696;
 Best Local Similarity 99.4%; Pred. No. 9.1e-128;
 Matches 488; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
 QY 296 CTCGCCCTGAGACATATTTTGGCATGTGTCTCAACAGATGTACCACTGTCCATAGAA 355
 1 CGTCTCCGACGACATATTTTGGCATGTGTCTCAACAGATGTACCACTGTCCATAGAA 60
 Db 356 TACTGAAAATGAGACAGATGCTGGCAGATGAGAAAGACCTAGTGGCAGCCCTGAGAGGT 415
 61 TACTGAAAATGAGACAGATGCTGGCAGATGAGAAAGACCTAGTGGCAGCCCTGAGAGGT 120
 QY 416 TTGCAGTTTTCCCAAGCTCAAGTGTGTTGGATCAGGGCAGCTGATGAACTGTGT 475
 121 TTGCAGTTTTCCCAAGCTCAAGTGTGTTGGATCAGGGCAGCTGATGAACTGTGT 180
 Db 476 TTGTGGAATGCACAGTCATACAAATTAATATAGATGTGTAGTGTAAAGTGGCTCCTT 535
 181 TTGTGGAATGCACAGTCATACAAATTAATATAGATGTGTAGTGTAAAGTGGCTCCTT 240
 QY 536 GGGGGCATGTGCATTTTCTCTAATGGAAGCTTTTGTGTCAGTGGCTCTCATGTGGTGA 595
 241 GGGGGCATGTGCATTTTCTCTAATGGAAGCTTTTGTGTCAGTGGCTCTCATGTGGTGA 300
 Db 596 TTAAACAGTGTGGGATGATAAATAGAGTGTCTGCATATGTAAGAAAGACATGATCTTGG 655
 301 TTAAACAGTGTGGGATGATAAATAGAGTGTCTGCATATGTAAGAAAGACATGATCTTGG 360
 QY 656 AATTACCTGCTGCAGATTTTCTTCTTCAAGCCAGTTTCTGATGAGAACAGAGTCTTCAGTT 715
 361 AATTACCTGCTGCAGATTTTCTTCTTCAAGCCAGTTTCTGATGAGAACAGAGTCTTCAGTT 420
 Db 716 TTTTGCACGTGCATCATGTGTGTGTCAGATTTGCCAAGTCAAAATTTGATTTTCTTTTAC 775
 421 TTTTGCACGTGCATCATGTGTGTGTCAGATTTGCCAAGTCAAAATTTGATTTTCTTTTAC 480
 QY 776 CCATATCTTAG 786
 Db 481 CCATATCTTAG 491

RESULT 12
 AK011391
 LOCUS
 DEFINITION Mus musculus 10 days embryo whole body cDNA, RIKEN full-length
 enriched library, clone:2610014f08 product:hypothetical SAM domain
 (sterile alpha motif)/Modified RING finger domain/G-protein beta
 WD-40 repeats containing protein, full insert sequence.
 ACCESSION AK011391
 VERSION AK011391.1 GI:12847483
 KEYWORDS HTC; CAP trapper.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 JOURNAL 99279253
 MEDLINE 10349636
 PUBMED 10349636
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 JOURNAL 20499374
 MEDLINE 11042159
 PUBMED 11042159
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Kono, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,

Qy 787 ----- 786
Db 851 AATATAAAGACACTAAGTGGGACCTGCCCCCTGTTCTGGCTGTCTTTTTCACATG 910
Qy 787 ----- 786
Db 911 ATGGAAGATGCTTGATCGGGGCTCAGTGAATAATCTGTCATCATATGATTCGACC 970
Qy 787 ----- 786
Db 971 CTCAGAGTGTGCTACACACGCTGACTCAGCATACAGGTATGTTAGCATTTGCGTTTG 1030
Qy 787 ----- 786
Db 1031 CACCCAACTCTTCTTACTGCTACTGCTCAATGACAAAGACATGTAACATTTGGCAGT 1090
Qy 787 ----- 825
Db 1091 TTGACCTGAAACACCTTGCCAGAGAAAGCATGAACCCGCTGAAACATTTGACTGAA 1150
Qy 826 ATTGCTCAGAGAGAGTGTCTCAACATGCTTTTGTGACAAAGATTTAAAGATTTGTTG 885
Db 1151 AATGCTCAGAGAGAGATGTCTCCGTTGGCTTCTGCTCAAGGCTTGAAGACCTCGTGC 1210
Qy 886 GTATTTCAAGATGATTAACATGATGAAAGAACTGTTGATCTTACAAAGAAAGTGC 945
Db 1211 GTATTTCAAGGCAACACATGATGAGAAAGAACTATGCTCAACAAAGAAAGTGC 1270
Qy 946 TGCGCTGATGATTTGAAATGAAATCTCTAGAGCTGCGTAAAGTGTGAGAAATG 1005
Db 1271 TGCGCTGATGATTTGAAATGAAATCTCTAGAGCTGCGTAAAGTGTGAGAAATG 1330
Qy 1006 AAGAGCTCAGAGCAAGGTTAATCTCTTTCTTCAAGAAATCTGATGAATTAATATGTC 1065
Db 1331 AAGAGCTCAGAGCAAGGTTAATCTCTTTCTTCAAGAAATCTGATGAATTAATATGTC 1390
Qy 1066 CAATTAACAGAGCACTTATGAAAGATCCGCTCATGCAATGAGGCTATTCATATGAA 1125
Db 1391 CAATTAACAGAGCACTTATGAAAGATCCGCTCATGCAATGAGGCTATTCATATGAA 1450
Qy 1126 AGGAAGCAATGAAATTTGATTCAGCAAAAGAAAGCAAGTCCCATGCAATCTG 1185
Db 1451 GAGGAAGCAATGAAATTTGATTCAGCAAAAGAAAGCAAGTCCCATGCAATCTG 1510
Qy 1186 TTCTTCTTCAAGCGTACTTACACCAATAGAGCTCTGAAATGCGCATCAATAGATGC 1245
Db 1511 CTCTCCCTTCACTGCTACTGACCCCAACAGACATGAAAGTGCATCAACGATGCG 1570
Qy 1246 TGGAGACACCAAAAGTAAA 1266
Db 1571 TGGAGACGACGAGAGAGTAA 1591

RESULT 13
CN407189 463 bp mRNA linear EST 16-MAY-2004
LOCUS 1700600054632 GRN_PPREHP Homo sapiens cDNA 5', mRNA sequence.
ACCESSION CN407189
VERSION CN407189.1 GI:47394734
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 463)
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J., Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandilam, R., Ledkovski, J. and Stanton, L.W.
TITLE Transcriptional characterization elucidates signaling networks that control human ES cell growth and differentiation
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)
COMMENT Contact: Brandenberger R

Regenerative Medicine
Geron Corporation
230 Constitution Drive, Menlo Park, CA 94025, USA
Tel: 650 473 8658
Fax: 650 473 7760
Email: rbrandenberger@geron.com
Insert Length: 463 Std Error: 0.00.
Location/Qualifiers
1..463
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/issue_type="embryonic stem cells, DMSO-treated H9 cell line"
/note="Toiigo dT primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

ORIGIN
Query Match 36.4%; Score 463; DB 7; Length 463;
Best Local Similarity 100.0%; Pred. No. 3,6e-121; Indels 0; Gaps 0;
Matches 463; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 16 CTGCGCGGACAGTACCGGACCGCCGCTGGGACCTTGAAGCGGATCCCGCGGCC 75
Db 1 CTGCGCGGACAGTACCGGACCGCCGCTGGGACCTTGAAGCGGATCCCGCGGCC 60
Qy 76 CGCTCTGACAGGCTGTTTCTTCAATTAAGAAACATGTGAAACATGATTCACATTA 135
Db 61 CGCTCTGACAGGCTGTTTCTTCAATTAAGAAACATGTGAAACATGATTCACATTA 120
Qy 136 CTGATCATGATGACAGATGTAACCTGTCGCTTCTCTTCCCTTGGCTACTGCT 195
Db 121 CTGATCATGATGACAGATGTAACCTGTCGCTTCTCTTCCCTTGGCTACTGCT 180
Qy 196 CTTGGAACAAACAAATTCGCTGATCGTTAGCTGATTAAGTCAATGCAATCTC 255
Db 181 CTTGGAACAAACAAATTCGCTGATCGTTAGCTGATTAAGTCAATGCAATCTC 240
Qy 256 CATTGAAGTTCAATACCTATGCTGTCAGCTGCTGCTTCTCCCTTCAAGACATATTT 315
Db 241 CATTGAAGTTCAATACCTATGCTGTCAGCTGCTGCTTCTCCCTTCAAGACATATTT 300
Qy 316 TGGCATGTGTTCAACAGATGTAACCTGCTCTATGGAATATCTGAATAATGACATGC 375
Db 301 TGGCATGTGTTCAACAGATGTAACCTGCTCTATGGAATATCTGAATAATGACATGC 360
Qy 376 TGGCAGATGTAACAGCTAGTGGCAGCCCTGAGAGGTTTGCCAGTTTCCCAAGCT 435
Db 361 TGGCAGATGTAACAGCTAGTGGCAGCCCTGAGAGGTTTGCCAGTTTCCCAAGCT 420
Qy 436 CCAAGTGTGTTGGATCAGGGGACAGCTGATGAACTGTGTTT 478
Db 421 CCAAGTGTGTTGGATCAGGGGACAGCTGATGAACTGTGTTT 463

RESULT 14
AL533462/c 967 bp mRNA linear EST 24-MAR-2004
LOCUS AL533462 Homo sapiens ADULT BRAIN Homo sapiens cDNA clone
DEFINITION CS00N004Y015 3-PRIME, mRNA sequence.
ACCESSION AL533462
VERSION AL533462.2 GI:31260543
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 967)
AUTHORS Li, W.B., Gruber, C., Jesse, J., and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)

COMMENT On Feb 13, 2001 this sequence replaced gi:12796955.

Contact: Genoscope
Genoscope - Centre National de Séquençage
2 rue Gaston Crémieux, CP 5706 - 91057 Evry cedex - FRANCE
Email: secref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five primers
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.

This sequence belongs to sequence cluster 1653.r
For more information about this cluster, see
<http://www.genoscope.cns.fr/cdna?b=CSD00N004&cc=CE08ND1&c=1653.r>

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .967 |

Query Match 33.98; Score 431.4; DB 1; Length 967;

"/note="Organ: brain; Vector: pcDNA3SPORT6; 1st strand cDNA was primed with a NotI-oligo (cm) primer. 1st prime end is enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pcDNA3SPORT 6 vector. Library was not normalized."

| | | | |
|-----------|--------|------|-----------------|
| RESULT 15 | | | |
| BQ935162 | | | |
| LOCUS | | | |
| BQ935162 | 910 bp | mRNA | linear |
| | | | EST 16-AUG-2002 |

| | |
|------------|--|
| DEFINITION | AEENCCOURT_8748888 NIH_MGC_130 Mus musculus cDNA clone IMAGE:6334422 |
| ACCESSION | 5', mRNA sequence. |
| VERSION | B0895162 |
| KEYWORDS | B0895162.1 GI:22287176 |
| SOURCE | EST. |
| ORGANISM | Mus musculus (house mouse) |
| | Mus musculus |

| | |
|-----------|---|
| REFERENCE | 1 (bases 1 to 910) |
| AUTHORS | NIH-MGC http://mgc.nci.nih.gov/ . |
| TITLE | National Institutes of Health, Mammalian Gene Collection (MGC) |
| JOURNAL | Unpublished (1999) |
| COMMENT | Contact: Robert Strausberg, Ph.D. |

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .910 |

Plate: L1AM13795 row: d column: 12
High quality sequence stop: 601.

ORIGIN

| | | | | |
|---------------------------|-------|---------------------|-----------|-------------|
| Query Match | 32.8% | Score 417.4; | DB 5; | Length 910; |
| Best Local Similarity | 77.9% | Pred. No. 5.5e-108; | | |
| Matches 529; Conservative | 0; | Mismatches 146; | Indels 4; | Gaps 2 |

/note="Origin: otcyabts; Vector: pCMV-SPORT6.1; Site_1:
 EcoRV; Site_2: NotI; Cloned unidirectionally. Primer:
 Otcy dt. Average insert size 1.95 kb. Constructed by
 Neogen, Invitrogen Corp. Note: this is a NIH MGC Library."

```

Db      420 TAGCTCATGTGCGCTGTGCGTTTCTCCGATGAGGCGCTTGTGTCACTGGCTCCTC 479
Qy      587 ATGTGTGATTTTAACAAGTGTGGGATGATTAATAATGAGGTGTCTGATAGTGAATAAGACA 646
Db      480 GGGCGGGAGCTTGACAGTGTGGATGACAGATGAGGTGTCTACACAGCGAAGGCGCA 539
Qy      647 TGAICTTGAATTACTGTGCTGCGATTTTCTTCAACGCCAGTTCTGATGAGAACAAAG 706
Db      540 CGATCTGGGATGACCTGCTGACGCTTTCTCTCACAGCCTCTCTGTGGCGAG--AAGG 596
Qy      707 TCTTCACTTTTTCGACTGCGATCATGTGTACGATTTGCCAAGTCAAAATTGATTTGT 766
Db      597 CTTCCAGTCTTACCAAGTTGGCGTCACTGTGTCAAGACTGTGAATCAAACTGTGGGCTGT 656
Qy      767 TTCTTTTACCCTATCTTA 785
Db      657 TACTATTACCGGTGTCTTA 675

```

Search completed: February 5, 2005, 08:18:30
 Job time : 4603 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 5, 2005, 10:22:21 ; Search time 5682 Seconds
(without alignments)
3274.693 Million cell updates/sec

Title: US-10-077-111-13

Perfect score: 2047
Sequence: 1 MYXLIHTLADHGDDVNCAP.....LTPNRTIKMAINRWLETHOK 384

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 segs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgml_1/USFTO.spool/US10077111/runat_04022005_121052_6802/app_query.fasta.1.553
-DB=GenEmbl -GEMT=faa1ap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blowum62 -TRANS=human40.cdi -LIST=45
-LOCALIZATION=200 -THR SCORE=pctc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USBR=US10077111.@CGM 1.1.4200.@runat_04022005_121052_6802 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb_ba:*
2: gb_ncg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_ets:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 2031 | 99.2 | 1564 | 9 | BC029520 Homo sapi |
| 2 | 1984 | 96.9 | 1996 | 6 | CQ491312 Sequence |
| 3 | 1984 | 96.9 | 1996 | 6 | CQ497179 Sequence |
| 4 | 1978 | 96.6 | 1811 | 6 | AX748164 Sequence |

| | | | | | | |
|----|--------|------|--------|----|----------|--------------------|
| 5 | 1978 | 96.6 | 1811 | 9 | AK093494 | AK093494 Homo sapi |
| 6 | 1978 | 96.6 | 1844 | 6 | AR339277 | AR339277 Sequence |
| 7 | 1978 | 96.6 | 1919 | 9 | AK129983 | AK129983 Homo sapi |
| 8 | 1823.5 | 89.1 | 1254 | 6 | C0729621 | C0729621 Sequence |
| 9 | 1617 | 79.0 | 2130 | 10 | BC061948 | BC061948 Mus muscu |
| 10 | 1464 | 71.5 | 3818 | 5 | AJ719461 | AJ719461 Gallus ga |
| 11 | 1373 | 67.1 | 1547 | 10 | BC050792 | BC050792 Mus muscu |
| 12 | 1226.5 | 59.9 | 1291 | 6 | AX076755 | AX076755 Sequence |
| 13 | 733 | 35.8 | 147021 | 2 | AC006501 | AC006501 Homo sapi |
| 14 | 731 | 35.7 | 245795 | 2 | AC096419 | AC096419 Rattus no |
| 15 | 731 | 35.7 | 315463 | 2 | AC122070 | AC122070 Rattus no |
| 16 | 728 | 35.6 | 177089 | 9 | AC008277 | AC008277 Homo sapi |
| 17 | 677 | 33.1 | 196403 | 2 | AC102767 | AC102767 Mus muscu |
| 18 | 677 | 33.1 | 209437 | 2 | AC111119 | AC111119 Mus muscu |
| 19 | 677 | 33.1 | 217352 | 10 | AL929242 | AL929242 Mouse DNA |
| 20 | 627.5 | 30.7 | 446 | 6 | C0473162 | C0473162 Sequence |
| 21 | 626.5 | 30.6 | 409 | 6 | C0482331 | C0482331 Sequence |
| 22 | 621.5 | 30.4 | 441 | 6 | C0503452 | C0503452 Sequence |
| 23 | 621.5 | 30.4 | 441 | 6 | C0512287 | C0512287 Sequence |
| 24 | 597 | 29.2 | 366 | 6 | C0096643 | C0096643 Sequence |
| 25 | 597 | 29.2 | 366 | 6 | C0135438 | C0135438 Sequence |
| 26 | 597 | 29.2 | 366 | 6 | C0218732 | C0218732 Sequence |
| 27 | 597 | 29.2 | 366 | 6 | C0294537 | C0294537 Sequence |
| 28 | 597 | 29.2 | 366 | 6 | C0331208 | C0331208 Sequence |
| 29 | 552 | 27.0 | 2279 | 3 | AK173854 | AK173854 Ciona int |
| 30 | 532 | 26.0 | 297 | 6 | C0109842 | C0109842 Sequence |
| 31 | 532 | 26.0 | 297 | 6 | C0148540 | C0148540 Sequence |
| 32 | 532 | 26.0 | 297 | 6 | C0231810 | C0231810 Sequence |
| 33 | 532 | 26.0 | 297 | 6 | C0307124 | C0307124 Sequence |
| 34 | 532 | 26.0 | 297 | 6 | C0344013 | C0344013 Sequence |
| 35 | 318 | 15.5 | 399 | 6 | AX071746 | AX071746 Sequence |
| 36 | 311 | 15.2 | 153866 | 9 | AC009307 | AC009307 Homo sapi |
| 37 | 308.5 | 15.1 | 466 | 6 | C0097537 | C0097537 Sequence |
| 38 | 308.5 | 15.1 | 466 | 6 | C0136375 | C0136375 Sequence |
| 39 | 308.5 | 15.1 | 466 | 6 | C0219679 | C0219679 Sequence |
| 40 | 308.5 | 15.1 | 466 | 6 | C0258151 | C0258151 Sequence |
| 41 | 308.5 | 15.1 | 466 | 6 | C0295845 | C0295845 Sequence |
| 42 | 308.5 | 15.1 | 466 | 6 | C0332079 | C0332079 Sequence |
| 43 | 300 | 14.7 | 180 | 6 | C0110706 | C0110706 Sequence |
| 44 | 300 | 14.7 | 180 | 6 | C0149443 | C0149443 Sequence |
| 45 | 300 | 14.7 | 180 | 6 | C0232723 | C0232723 Sequence |

ALIGNMENTS

| | | | | | |
|------------|--|--|------|--------|-----------------|
| RESULT 1 | BC029520 | 1564 bp | mrna | linear | PRI 20-JUL-2004 |
| LOCUS | BC029520 | Homo sapiens WD repeat and SAM domain containing 1, mRNA (cDNA | | | |
| DEFINITION | clone MGC:33855 IMAGE:5301559), complete cds. | | | | |
| ACCESSION | BC029520.1 | GI:20810486 | | | |
| VERSION | BC029520 | | | | |
| KEYWORDS | MGC. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 1564) Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Shuter,G.D., Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Altschul,S.P., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Usetin,T.B., Toshlyuk,S., Schenck,T.E., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.W., Gay,L.J., Huily,S.W., Villalón,D.R., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shcherchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., | | | | |

| | | | |
|------------|---|---|-----------------|
| Db | 1226 | CCTTAGCGGCTACTTACACCAATTCGACTCTGAAATGGCCATCAATGATGGCTGGAG | 1285 |
| QY | 381 | ThrhIsgInLYs 384 | |
| Db | 1286 | ACACACCAAAAG | 1297 |
| RESULT 2 | | | |
| LOCUS | CQ491312 | 1996 bp | DNA |
| DEFINITION | Sequence 23179 from Patent WO0160860. | linear | PAT 30-JAN-2004 |
| ACCESSION | CQ491312 | | |
| VERSION | CQ491312.1 | GI:41456931 | |
| KEYWORDS | | | |
| SOURCE | | | |
| ORGANISM | Homo sapiens (human) | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | |
| REFERENCE | Schlegel, R., Endege, W.O. and Monahan, J.E. | | |
| AUTHORS | 1 | | |
| TITLE | Genes differentially expressed in human prostate cancer and their use | | |
| | Patent: WO 0160860-A 23179, 23-AUG-2001, Millennium Predictive Medicine, Inc. (US) | | |
| JOURNAL | | | |
| FEATURES | | | |
| source | 1..1996 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="unassigned DNA" | | |
| | /db_xref="taxon:9606" | | |
| ORIGIN | | | |

| | |
|------------------------|-----------|
| Alignment Scores: | |
| Pred. No.: | 1,96e-172 |
| Score: | 1984.00 |
| Percent Similarity: | 80.46% |
| Best Local Similarity: | 80.46% |
| Query Match: | 96.92% |
| DB: | 6 |
| | Gaps: 1 |

| | | | |
|----|-----|--|-----|
| OY | 1 | MetValysLeuIleHsTrpLeuAlaAspHisGlyAspAspValAsnCysGlyAlaPhe | 20 |
| Db | 16 | ATGTTAAACTGATTACACATTAGCTGATCAGTGAAGATGACATGCTGTCCTTC | 75 |
| OY | 21 | SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTyrSerLeuArg | 40 |
| Db | 76 | TCCCTTTCCCTCTGGCTACTCTCTCTGGACAAACCAATTGGCTGACTGTTACGT | 139 |
| OY | 41 | AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyrAlaValHisCysCys | 60 |
| Db | 136 | GACTTAACTGAATGCCACATTTCCATTGAAGTTTCATACCTAATGCTGCATGCTGC | 199 |
| OY | 61 | CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu | 80 |
| Db | 196 | TGTTTCCCTCCCTCAGACATATTTTGGCATCGTGTGTCAACAGATGGTACCATCTCTTA | 259 |
| OY | 81 | TrpAsnThrGluAsnGlyIleMetLeuAlaValMetGluGlnProSerGlySerProVal | 100 |
| Db | 256 | TGGAAATCTAATAAATGACAGATGCTGGCAGTGATGGAACACCTTAATGGCAGCCCTGTG | 319 |
| OY | 101 | ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyValAlaAspGlyThr | 120 |
| Db | 316 | AGGGTTTGGCAGTTTCCCAAGCTCACGTGTTTGGCATCAGGGCAGCTGATGAACT | 379 |
| OY | 121 | ValValLeuTrpAsnAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysAspGly | 140 |
| Db | 376 | GTGGTTTGTGGAATGCACAGTCATCAAAATTATTATAGTGTGTATGTTTAAAGATGCC | 439 |
| OY | 141 | SerLeuAlaIaCysAlaPheSerProAsnGlySerPhePheValThrGlySerCys | 160 |
| Db | 436 | TCCTTGGCGGATGTGATTTTCTCTTAAGGAAGCTTTGTGTACATGGCTCTTCACAGT | 499 |

| | | | |
|----|------|---|------|
| QY | 161 | GLYAspLeuThrValITTPAspAspLysMetArgCysLeuHisSerGluValValHisAsp | 180 |
| Db | 496 | GStGATTTAAcAGTGTGGATGATPAAATGAAGTGTCTGCAATAGTGAATAAAAGCCACATAT | 555 |
| QY | 181 | LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu | 200 |
| Db | 556 | CTTGGAATTAACtCTGCTGGATTTTCTTCTTCAAGCAGtTTCTGATGGAGAAcAAAGCTT | 615 |
| QY | 201 | GlnPhePheArgLeuValAspCysGlyGlnAspCysGlnValIleTyrIleValSer | 220 |
| Db | 616 | CAGTTTTTCCAGCTGGCATCATGTGTGTCAGAGATTGCCAAcGTAATAATTGGATTGTTTCT | 675 |
| QY | 221 | PheThrHisIleLeu----- | 225 |
| Db | 676 | TTTACCCATATCTTAGCTTTGAAATTAAATAATATAAAAGTACAGTAgTGGCACtTGCT | 735 |
| QY | 225 | ----- | 225 |
| Db | 736 | CGTGTTCGGCTTGTGTCTTTTCCAGATGtGGGAGATGCTAGTCTAGGCTCAGTGGAT | 795 |
| QY | 225 | ----- | 225 |
| Db | 796 | AAGTCTGTCATAGTATATGATACTAATACTGAGATATATCTTCAcACATGTACTCAGCAC | 855 |
| QY | 225 | ----- | 225 |
| Db | 856 | ACCAAGTATGTCAcAACTTGCTTTGGCACCTTAATACCTTTTACTTGCTTACTGTGTCA | 915 |
| QY | 226 | -----AlaArg | 228 |
| Db | 916 | ATGGAcAAAcAGAGAAcGATGTGGCAATTGACCTGGAAAcACTTGGCAAGAGGCGC | 975 |
| QY | 229 | ThrGluHisGlnLeuLysGlnPheThrGluAspTyrSerGluGluValValSerThrTyr | 248 |
| Db | 976 | ACAGAAcATcAGCTGAAGCAATTATACCGAAGATTGTGCAGAGAGAGATGTCTCAAcATGG | 1035 |
| QY | 249 | LeuCysValIaglnAspLeuLysAspLeuValGlyIlePheLysMetAsnAsnIleAspGly | 268 |
| Db | 1036 | CTTTGTGCACAGATTATAAAAGATCTGTGTGTGATTTTTCAGAGATGAATTAACATTGATGGA | 1095 |
| QY | 269 | LysGluLeuLeuAsnLeuThrLysGlnSerLeuIleAspAspLeuLysIleGlnSerLeu | 288 |
| Db | 1096 | AAAGAcCTGTGAATCTTACAAAGAAAGAcCTGGCTGATGATTTTGAAATTTGAATCTCTA | 1155 |
| QY | 289 | GlyLeuArgSerLysValLeuArgLysIleGlnGluLeuArgThrLysValLysSerLeu | 308 |
| Db | 1156 | GGACTGCTACTATAAGTGTGAGGAAATTTGAAGAcCTCAGGACCAAGGTTAAATCCCTT | 1215 |
| QY | 309 | SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro | 328 |
| Db | 1216 | TCTTCAGGAATTCCTGATGATTTATATATGTCCATTAATCTAGAGAACTTATATAAGATTCG | 1275 |
| QY | 329 | ValIleAlaSerAspGlyTyrSerTyrGluLysGlnIleMetGluAsnTyrIleSerLys | 348 |
| Db | 1276 | GTCATCGCATCAGATGGCTATTCTATATGAAGAAAGAGAAATTTGGATTCAGCAA | 1335 |
| QY | 349 | LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn | 368 |
| Db | 1336 | AAGAAGAcCTACAGATGCCATACAAAcCTTGTTCTTCTTCAGGCGGTACTTACACCAAT | 1395 |
| QY | 369 | ArgThrLeuLysMetAlaIleAsnArgTyrPheGlnThrHisGlnLys | 384 |
| Db | 1396 | AGAcATCTGAATAATGGCCATCATATAGATGGCTGGAGACACCCAAAG | 1443 |

| RESULT 3 | | | | |
|------------|---------------------------------------|-------------|-----|-------------------------|
| LOCUS | C0497179 | | | |
| DEFINITION | Sequence 29046 from Parent W00160860. | 1996 bp | DNA | |
| ACCESSION | C0497179 | | | |
| VERSION | C0497179.1 | GI:41462815 | | |
| KEYWORDS | | | | |
| SOURCE | ‘Homo sapiens (human) | | | |
| ORGANISM | Homo sapiens | | | |
| | | | | Linear PAT 30--JAN-2004 |

REFERENCE 1
AUTHORS Schlegel, R., Endege, W.O. and Monahan, J.E.
TITLE Genes differentially expressed in human prostate cancer and their use
JOURNAL Patent: WO 0160860-A 29046 23-AUG-2001;
Millennium Predictive Medicine, Inc. (US)
FEATURES
source 1. .1996
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 1.96e-172 Length: 1996
Score: 1984.00 Matches: 383
Percent Similarity: 80.46% Conservative: 0
Best Local Similarity: 80.46% Mismatches: 1
Query Match: 96.92% Indels: 92
DB: Gaps: 6
US-10-077-111-13 (1-384) x CQ497179 (1-1996)
QY 1 MetValIysLeuIleHisThrLeuAlaAspHisGlyAspAspValaAsnCySValaPhe 20
DB 16 ATGGTGAACCTGATTCACACATTAGCTGATCATGTCAGCATGTCACCTGCGCTTC 75
QY 21 SerPheSerLeuLeuAlaThrCySerSerLeuAspIleThrIleAspGluTyrSerLeuArg 40
DB 76 TCCCTTTCCTCTGGCTACTTGTCTCTTGGACAAACAAATTCCTGCTACTGCTTAAGT 135
QY 41 AspPheThrGluLeuProHisSerProLeuLeuPheHisThrTyrAlaValHisCySVala 60
DB 136 GACTTACTGAACTGCGCACATTCCTCATTAAGTTTCACTACTGCTGCTGCTGCTGC 195
QY 61 CyasPheSerProSerGlyHisIleLeuAlaSerCySValaThrAspGlyThrThrValLeu 80
DB 196 TGTTTCCCTCCCTCAGAGACATATTTGGCATCGGTTCACAGATGGTACCATGCTCTA 255
QY 81 ThrAsnThrGluLeuAspGlyIleMetLeuAlaValMetGluIleProSerGlySerProVal 100
DB 256 TGGAAATCTGAATAATGACAGATGCTGAGTGTGAGAACAGCCTACTGCGAGCCCTGTG 315
QY 101 ArgValCySglnPheSerProAspSerThrCySLeuAlaSerGlyAlaIleAspGlyThr 120
DB 316 AGGCTTTCAGAGTTTCCCAAGACTCCAGCTGTTGGCATCAGGAGGAGCTGATGAACT 375
QY 121 ValValLeuThrAspAlaGlnSerTyrLysLeuTyrArgCySglnSerValIlyAspGly 140
DB 376 GTGGTTTGTGAAATGACAGTCAATCAATTAATATGATGTGTTAAAGATGGC 435
QY 141 SerLeuAlaIleCySValaPheSerProAspGlySerPhePheValThrGlySerSerCyS 160
DB 436 TCCCTGGCGGCACTGTGATTTTCTCCPAAATGGAAGCTTCTTGTGCTGCTGCTCATGT 495
QY 161 GlyAspLeuThrValIleTPAspAspIleMetArgCySValaHisSerGlyIleValaHisAsp 180
DB 496 GGTGATTTAAACAGTGTGGATGATTAATAAGGTGTGCACTGATGAAAAAGCACAAGAT 555
QY 181 LeuGlyIleThrCySAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
DB 556 CTGGGAATTAACCTGCTCGATTTTCTTCAACGCACTTCTGATGGAAGAACAGGCTT 615
QY 201 GlnPhePheArgLeuAlaSerCySglnGlnAspCySglnValIlyIleTPIleValSer 220
DB 616 CAGTTTTTTTGACTGGCATCATGTGTGAGGATTCAGGATTCAGCAAAATTTGGATTGTTCT 675
QY 221 PheThrHisIleLeu----- 225
DB 676 TTTAACCAATCTTAAGCTTTGAATTAAATATATAAGTACATGAGTGGGCACTGTGCT 735

QY 225 ----- 225
DB 736 CCTGTTCTGCTGTGCTTTTCCATGATGGCAGATGCTAGTCAAGGTGATGAT 795
QY 225 ----- 225
DB 796 AAGTCTGTCACTAGTATATGATTAATTAAGTGAATATGATTAATGATGATGATGAT 855
QY 225 ----- 225
DB 856 ACCAGGATGTCAACATGTTGCTTTTGACACCTAATACCTTTTACTGCTAGTGTCA 915
QY 226 -----11aa-garg 228
DB 916 ATGACAAAACAGTGAACATCTGCAATTGACCTGAAAACACTTTGCCAAGCAAGCGCC 975
QY 229 ThrGluHisGlnLeuLeuGlnPheThrGluAspThrSerGlnGluValIleSerThrTrp 248
DB 976 ACAGAACATCAGCTGAAGCAATTTACGAAAGATTGTCAGAGAGGATGTCTCAACATGG 1035
QY 249 LeuCyAlaGlnAspLeuIlyAspLeuValGlyIlePheLysMetAsnAsnIleAspGly 268
DB 1036 CTTTGTGCACAAATTTAAAGATCTTGTGATTTTCAAGATGAATACATGATGAGA 1095
QY 269 LysGluLeuLeuAsnLeuThrIlyGlnSerLeuAlaAspAspLeuIlyGlnSerLeu 288
DB 1096 AAAGAACTGTTGAATCTTCAAAAGAAAGCTGGCTGATGATTTGAAATCTCTA 1155
QY 289 GlyLeuAspSerIlyValLeuArgIlyIleGluIleLeuArgThrIlyValIlySerLeu 308
DB 1156 GAGCTGCGTGAAGTGAAGCTGAGAAATTTGAAGACTCAGGACCAAGGTTAAATCCCT 1215
QY 309 SerSerGlyIleProAspGluPheIleCySProIleThrArgGluLeuMetLysAspPro 328
DB 1216 TCTTCAGGAATCTCTGATGATTAATATGATCCAAATACAGAACTTATGAAAGATCCG 1275
QY 329 ValIleAlaSerAspGlyTyrSerTyrGluLysGluAlaMetGluLeuThrIleSerLys 348
DB 1276 GTCATGCAATCAGATGCTATTCATATGAAAGAAAGCAATGGAATTTGATCACCAAA 1335
QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerLysValLeuThrProAsn 368
DB 1336 AAGAAACGTACAACTCCATGACAAATCTTGTCTTTCAGCGGTACTTACACCAAAAT 1395
QY 369 ArgThrLeuLysMetAlaIleAsnArgTyrLeuGluThrHisGlnLys 384
DB 1396 AGGACTCTGAATAATGGCCATCAATATGATGCTGGAACACACCAAAAG 1443
RESULT 4
AX748164 1811 bp mRNA linear PAT 20-JUN-2003
LOCUS Sequence 1689 from Patent EP1308459.
DEFINITION AX748164
ACCESSION AX748164
VERSION AX748164.1 GI:32132552
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Isozaki, T., Sugiyama, T., Otsuki, T., Wakamatsu, A., Sato, H., Ishii, S., Yamamoto, J.I., Isono, Y., Hio, Y., Otsuka, K., Nagai, K., Irie, R., Tamechika, I., Seki, N., Yoshikawa, T., Otsuka, M., Nagahashi, K. and Masuno, Y.
TITLE Full-length cDNA sequences
JOURNAL Patent: EP 1308459-A 1689 07-MAY-2003;
Helix Research Institute (JP) ; Research Association for Biotechnology (JP)
FEATURES
source 1. .1811
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

ORIGIN

Alignment Scores:

Pred. No.: 6,13e-172 Length: 1811
 Score: 1978.00 Matches: 382
 Percent Similarity: 80.25% Conservative: 0
 Best Local Similarity: 80.25% Mismatch: 2
 Query Match: 96.63% Indels: 92
 DB: 6 Gaps: 1

US-10-077-111-13 (1-384) x AK748164 (1-1811)

QY 1 MetValylsLeuIleHisThrLeuAlaAspHisGlyAspAspValaIasnCysCysAlaPhe 20
 Db 150 ATGCTGAACGATTCACACATTCATGATCATGTCAGATGTCACATGCTGAGCCCTTC 209
 QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuYrSerLeuArg 40
 Db 210 TCCCTTTCCTCTGGCTACTTGTCTCTGGACAAACAAATTCCTGTAAGTCTTACGT 269
 QY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrYrAlaValHisCysCys 60
 Db 270 GACTTACTGAACTGACCAACATTCATGAAATTCATGAACTGATGCTGCTGCTGCTG 329
 QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
 Db 330 TGTTCCTCCCTTCAGGACATATTTGGCATCGTGTTCACAGATGCTACCATCTGCTCTA 389
 QY 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
 Db 390 TGGAAATCTGAATAATGACAGATCTGCGAGTATGAAACAGCTTATGCGAGCCCTGTG 449
 QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
 Db 450 AGGCTTGGCCAGTTCCTCCAGACTCCAGTGTTCGATCAAGGCGACGCTGATGAACT 509
 QY 121 ValValLeuThrPheAsnAlaGlnSerYrIleLeuYrArgCysGlySerValLysAspGly 140
 Db 510 GTGCTTTTGTGAGATGACACATCTACAAATTAATAGATGTGATGTTAAAGATGCG 569
 QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPheValIleArgIleSerSerCys 160
 Db 570 TCCCTGGCGGATGTGATTTCTCTTAATGAAAGCTTCTTGTGACATGCTCTCTCATGT 629
 QY 161 GlyAspLeuThrValIleTrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
 Db 630 GGTGATTTAACAGTGTGGATGATMAATGAGGTGTCTGCAATGATGAAAGCAACATGAT 689
 QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
 Db 690 CTTGGAAATACCTCTGCGATTTTCTTCCACAGCCAGTTTCTGATGGAACAAAGTCTT 749
 QY 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer 220
 Db 750 CAGTTTGTGACATGCGATCATGTGTCAGATTGCCAAGCAAAATTTGGATTGTTCTT 809
 QY 221 PheThrHisIleLeu----- 225
 Db 810 TTTTACCATATCTTAAAGTTTGAATTAATAATATAAAGTACACAGTGGCGACATGCT 869
 QY 225 ----- 225
 Db 870 CCTGTTCTGCTGTGCTTTTCCCATGATGCGAGATGCTAGTCTCAGGCTCAGTGAT 929
 QY 225 ----- 225
 Db 930 AAGCTGTGATATGATATGATATGATATGATATGATATGATATGATATGATATGATATG 989
 QY 225 ----- 225
 Db 990 ACCAGTATGTCAACAATTTGTCTTTGGCACTTAATACCTTTTACTTGTACTGCTTCA 1049
 QY 226 -----AlaArgArg 228

Db 1050 ATGACAAACACATGACATCTGCGCAATTGACCTGGAAACACTTGGCCAAAGAGAC 1109
 QY 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGlyGluGluValIleSerThrTrp 248
 Db 1110 ACAGAACATACAGCTGAAAGCAATTTTCCAGAGATTTGTCAAGAGAGATGTCACATGCG 1169
 QY 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnAlaIleAspGly 268
 Db 1170 CTTTGTGCAACAAGTTTAAAGATCTTGTGTGATATTTCAAGATGATTAACATGATGAG 1229
 QY 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
 Db 1230 AAAGAACTGTTGAATCTTCAAAAGAAAGCTGCTGATGATTTGAAAATTTGAATCTCTA 1289
 QY 289 GlyLeuAspSerLysValIleuArgLysIleGluGlnLeuLysArgThrLysValLysSerLeu 308
 Db 1290 GGACTGCGTGTAAAGCTGAGAGAAATTTGAAGAGCTCAGACCAAGTTTAAATCCCTT 1349
 QY 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
 Db 1350 TCTTCAGAAATTCCTGATGATTTATATGTCCTAATACAGAACTTAAGAAAGATCCG 1409
 QY 329 ValIleAlaSerAspGlyYrSerYrGluLysGluAlaMetGluAsnTrpIleSerLys 348
 Db 1410 GTATGTCATCAGATGCTGCTATTCATATGAAAGAAAGCAATGAAATTTGATCAGCAAA 1469
 QY 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
 Db 1470 AAGAAAGTCAAGATCCCAAGCAACAAATCTTCTTCCACGCGTACTTACACCAAT 1529
 QY 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGlnThrHisGlnLys 384
 Db 1530 AGGACTGTGAATATGCGCATCAATAGATGCTGAGACACACCAAAAG 1577

RESULT 5

AK093494

LOCUS

DEFINITION

Homo sapiens cDNA FLJ36175 fis, clone T88712026453, weakly similar to PUTATIVE SERINE/THREONINE-PROTEIN KINASE PKMA (EC 2.7.1.-).

AK093494.1 GI:21752384

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 Ota, T., Suzuki, Y., Nishikawa, T., Otsuki, T., Sugiyama, T., Irie, R., Makamatsu, A., Hayashi, K., Sato, H., Nagai, K., Kimura, K., Makita, H., Sekine, M., Ohtsuka, M., Nishi, T., Shidohara, T., Tanaka, T., Ishii, S., Yamamoto, J., Saito, K., Kawai, Y., Isono, Y., Nakamura, Y., Nagahara, K., Murakami, K., Yasuda, T., Iwayanagi, T., Wagatsuma, M., Shiratori, A., Sudo, H., Hosoiri, T., Kaku, Y., Kodaira, H., Kondo, H., Sugawara, M., Takahashi, M., Kanda, K., Yokoi, T., Furuya, T., Kikawa, E., Omura, Y., Abe, K., Kamihara, K., Katsura, N., Sato, K., Tanikawa, M., Yamazaki, M., Ninomiya, K., Ishibashi, T., Yamashita, H., Murakawa, K., Fujimori, K., Tanai, H., Kimata, M., Watanabe, M., Hiraoaka, S., Chiba, Y., Ishida, S., Ono, Y., Takiguchi, S., Watanabe, S., Yoshida, M., Houta, T., Kusano, J., Kanehori, K., Takahashi, Fujii, A., Hara, H., Tanase, T., Nomura, Y., Togaya, S., Komai, F., Hara, R., Takeuchi, K., Arita, M., Imose, N., Kusashino, K., Yuki, H., Oshima, A., Sasaki, N., Aotsuma, S., Yoshikawa, Y., Matsunawa, H., Ichihara, T., Shiohata, N., Sano, S., Moriya, S., Komiyama, H., Satoh, N., Takami, S., Terashima, Y., Suzuki, O., Nakagawa, S., Senoh, A., Mizoguchi, H., Goto, Y., Shimizu, F., Wakebe, H., Hishigaki, H., Watanabe, T., Sugiyama, A., Takemoto, M., Kawakami, B., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Tashiro, H., Tanigami, A., Fujizawa, T., Ono, T., Yanada, K., Fujii, Y., Ozaki, K., Hirao, M., Omori, Y., Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y., Okamoto, S., Okitani, R., Kawakami, T., Noguchi, S., Itoh, T., Shigeta, K., Senba, T., Matsunura, K., Nakajima, Y., Mizuno, T., Morinaga, M., Sasaki, M., Togashi, T.,

TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
Oyama, M., Hata, H., Watanabe, M., Komatsu, T., Mizushima-Sugano, J., Sato, T., Shirai, Y., Takahashi, Y., Nakagawa, K., Okumura, K., Nagase, T., Nomura, N., Kikuchi, H., Masuo, Y., Yamashita, R., Nakai, K., Yada, T., Nakamura, Y., Ohara, O., Isegai, T. and Sugano, S.
Complete sequencing and characterization of 21,243 full-length human cDNAs
Nat. Genet. 36 (1), 40-45 (2004)
14702039

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
Unpublished
3 (bases 1 to 1811)
Isegai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isegai, FLJ Project (HRI Team); 2-6-7 Kazuwa-Kamatari, Kibaraazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB) / cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center; National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
Location/Qualifiers
1. 1811

FEATURES
source
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="FESTI2026453"
/issue_type="test1a"
/clone_lib="FESTI2"
/note="Cloning vector: pME18SFL3"
150..1580
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC04184.1"
/db_xref="GI:21752385"
/translation="MVKLIHTLADHDGDVNCASFPSLTLATGSLDKTIRLYLRDFTL
LPHSPKFTYAVHCCFSPSGHILASCTDGTIVLMTENGMLVMEQPSGSPYRV
CQSPSTCIASGADGTVYLMNAQSYKLYRCGSVDGSLAACAFSPNGSPVYTGSSC
GLTVMDDKMRGCHSEKHAHDLGITCCFSGSPQVSDGQIGFRIASCGDDCVKXIM
VSFTHLIDPELKYKRTLSGCHABYLACASPHDQMLVSGSDVSIVYDPTNLIHT
LTOHTRYVTTCAFPNTLLATSGMDKTIVIKQFDLTLCQASTHQLQFTEDISE
EDVSTWCAQDLKDLVIGIFPMNNIDGELNLTKESLADLKLKESGLRSKYLAKIE
LRTKVSLSGSGIDPEIFCIPTREIMKDPVLASDGYSEKEMENWISKKRKTSPTMNL
VLPSAVLTPERTLKMAINRWLETHOK"

CDS

ORIGIN
Alignment Scores:
Pred. No.: 6,13e-172 Length: 1811
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 9 Gaps: 1
US-10-077-111-13 (1-384) x AK093494 (1-1811)

QY 1 MewVallyleuileuilehsthrleuAlaAspHieGlyAspAspValaIaenCyseCyalaIaphe 20
DB 150 ATGGTGAACCTGATTCACACATTAGCTGATCATGTGAGCATGTCAACTGCTGTGCTTC 209

QY 21 SerPheSerleuLeuAlaThrCyseSerleuAspIlyThrIleArgIeuIyrSerleuArg 40
DB 210 TCGTTTTCCTCTTGGCTCTGCTCTTGGACAAAAACAATTCGCTGACTGCTTACGT 269
QY 41 AspPheThrGlyIeuProHisSerProLeuIyPheIsthrTyraIaValHisCyseCy 60
DB 270 GACTTACTGAACTGGCACATTTCTCAATTGAAGTTTCATACCTATCTGTCCCTGCTG 329
QY 61 CyseSerProSerGlyHisIleleuAlaSerCyseSerThrAspGlyThrThraIleu 80
DB 330 TGGTTTCCCTCCCTCAGACATATTTTGGCATGCTGTTCACAGATGTGACACTGCTCA 389
QY 81 TrpAsnThrGlyIuAsnGlyIuMetleuAlaValMetGluGlnProSerGlySerProVal 100
DB 390 TGGAAATACGAAATATGACAGATGCTGGCAGTGTGAACAGCTTGTGTGCGACGCTGTG 449
QY 101 ArgValCyGluPheSerProAspSerThrCyseleuAlaSerGlyIaAlaAspGlyThr 120
DB 450 AGGGTTTGGCAGTTTCCCGACGCTCAGCTGTGGCATCAGGGGACGCTGATGGAGCT 509
QY 121 ValValleuThrAsnAlaGlnSerTyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIyIy 140
DB 510 GTGGTTTGTGGAAATGACAGTCAATACAAATTAATAGATGTGTATGAAGATGGC 569
QY 141 SerleuAlaAlaCyalaPheSerProAsnGlySerPhePheValThrGlySerSerCy 160
DB 570 TCGTTGGCGGACATGTCATTTTCTCTTAATGGAAGCTTCTTGTGCTAGCTGCTCTCATGT 629
QY 161 GlyAspLeuThrValTrpAspAspIyMetIaCyseIuHisSerGlyIyIyIyIyIyIyIy 180
DB 630 GGTGATTTAAAGTGTGGATGATTAATAAGAGTGTCTCATATGTAAAGAACATGAT 689
QY 181 LeuGlyIleThrCyseCyAspPheSerSerGlnProValSerAspGlyGluGlnGlyIeu 200
DB 690 CTGGAAATTAACGTCGCGCATTTTCTTTCACAGCCAGTTTCTGATGGAACAAGCTCTT 749
QY 201 GlnPhePheArgleuAlaSerCyGlyGlnAspCyseGlnValIyIleTrpIleValSer 220
DB 750 CAGTTTTCCTGACTGCGCATCATGTGTGTCAGATTCGCAAGTCAAAATTTGGATTGTTCT 809
QY 221 PheThrHisIleIeu----- 225
DB 810 TTTAACCAATATCTAGCTTTTGAATTAATAATTAACACAGTACAGTGGCAGCTGTGCT 869
QY 225 ----- 225
DB 870 CCTGTTCTGGCTTGTGCTTTTCCCATGATGGGACAGATGCTAGCTCAGAGTCAAGTAT 929
QY 225 ----- 225
DB 930 AAGTCTGATAGTATATGATACTAATACTGAGAATATACTTCACACATTGACTCAGAC 989
QY 225 ----- 225
DB 990 ACCAGGATATGCACAATTGTGCTTTTGCACCTAATACCTTTTACTTCTACTGCTTCA 1049
QY 226 -----11AArgArg 228
DB 1050 ATGGACAAAACAGTGAACATCTGGCAATTGACCTGGAAAACCTTTGCCAAGCAGAGACC 1109
QY 229 ThrGluHisGlnIleuIyGlnPheThrGluAspTrpSerGluGluValIaIaSerThrTrp 248
DB 1110 AACGAACATCAGCTGAAGCAATTTACCGAAGATTTGCTCAGAGAGAGATCTTCAACATGG 1169
QY 249 LeuCyAlaGlnAspIeuIyAspIeuValGlyIlePheIyMetCAsnAsnIleAspGly 268
DB 1170 CTTTGTGCAACAAGATTTAAAGATCTTGTGTGATTTTCMAAGATGAATTAACATTGATGGA 1229
QY 269 IyGluIleuIyAsnIeuThrIyGluIySerleuAlaAspAspIeuIyIleGluIySerIeu 288
DB 1230 AAAGAACGTGTGAATTTTCAAAAGAAAGTCTGGCTGAGATTTGAAAATTTGAATCTCTTA 1289
QY 289 GlyIeuArgSerIyValIeuArgIyIleGluGluIleuArgThrIyValIySerIeu 308

Db 1290 GGACTGGGTAAAGCTGAGAGAAATTCAGAGCTCAGACCAAGGTTAAATCCCTT 1349
Qy 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
Db 1350 TCTTCAGGAATTCCTGATGAAATTTAATGTCCTCAATACCTAGAGAACTTATGAAAGATCCG 1409
Qy 329 ValIleAlaSerAspGlyIleProSerIleGluGluAlaMetGluAsnTrpIleSerLys 348
Db 1410 GTCATGCCATCAGATGGCTATTCATATGAAAGAGCAATGAAATTTGGATCAGCAAA 1469
Qy 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db 1470 AAGAAAGTACAGATCCCATACAAATCTGTTCTTCACGGGTAATTACACCAAAAT 1529
Qy 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGluLys 384
Db 1530 AGGACTCTGAAATGGCCATCAATAGATGGCTGGAGACACCAAAAG 1577

RESULT 6

LOCUS AR339277 1844 bp DNA linear PAT 17-AUG-2003
DEFINITION Sequence 768 from patent US 6569662.
ACCESSION AR339277
VERSION AR339277.1 GI:33726134
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.

REFERENCE

1 (bases 1 to 1844)
Tang, Y.T., Zhou, P. and Drmanac, R.T.
Nucleic acids and polypeptides
JOURNAL Patent: US6569662/768 27-MAY-2003;
FEATURES
Source Location/Qualifiers
1..1844
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

Pred. No.: 6.28e-172 Length: 1844
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: 6 Gaps: 1

US-10-077-111-13 (1-384) x AR339277 (1-1844)

Qy 1 MetValLysLeuIleHisThrLeuAlaAspHisGlyAspAspValaAsnCysCysAlaPhe 20
Db 164 ATGGATGAACCTGATTCACACATAGCTGATCATGATGACATGTCACACTGCTGCTTC 223
Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuIleSerLeuArg 40
Db 224 TCCCTTTCCCTCTGGCTACTGCTGCTGTCGACAAAACAATTCGCTGACTGCTTACGT 283
Qy 41 AspheThrGluLeuProHisSerProLeuLysPheHisThrIleAlaValHisCysCys 60
Db 284 GACTTTCCTGAACTGCCACATTCCTCAATGAAATTTCAATACCTATGCTGCTGCTGCTGC 343
Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
Db 344 TGTTTCTCCCTTCAGACATATTTGGCATCGTGTTCACAGATGGTACCACCTGCTTA 403
Qy 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 404 TGGAAATCTGAAATGGACAGATCTGCGACATGAGAACAGCCTAGTGGACGCTGTG 463
Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
Db 464 AGGGTTTGCCAGTTTCCCGACACTCCAGCTGTTTGGCATCAGGGGCGAGCTGATGAACT 523

Qy 121 ValValLeuTrpAsnAlaGlnSerIleLysLeuIleThrArgCysGlySerValLysAspGly 140
Db 524 GTGGTTTGTGAATGACAGCTCATCAAAATTTATATAGATGTGTGTTAAAGATGCG 583
Qy 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
Db 584 TCTTGCGGCGCATGTGATTTCTCTAAAGAGCTCTTGTCACTGCTCTCAATGT 643
Qy 161 GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGluLysAlaHisAsp 180
Db 644 GGTGATTTAAACATGTGGATGATTAATAATAGGTGTCTGATAGTAAAGACATGAT 703
Qy 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlnGlyLeu 200
Db 704 CTGGAAATTAACCTGCTGCGATTTTCTTCAACGCCAGTTTCTGATGAGAAACAAGCTT 763
Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnAspCysGlnValLysIleTrpIleValSer 220
Db 764 CAGTTTCTTGCATGCGCATGTGTGACAGATTCGCAAGTTCAAAAATTTGGATTTGTTCT 823
Qy 221 PheThrHisIleLeu----- 225
Db 824 TTTAACCATATCTTAGGTTTGAATTAATAATATAAGTACACTGAGTGGCACTGTCT 883
Qy 225 ----- 225
Db 884 CTTGTTCTGCTGTGTGCTTTTCCATGATGGCAGATGCTAGTCAAGGTCATGTGAT 943
Qy 225 ----- 225
Db 944 AAGCTGTCTATGAT 1003
Qy 225 ----- 225
Db 1004 ACCAGTATGTCACAACTGCTGCTTTGCACTTAATACCTTTTACTGCTACTGCTCA 1063
Qy 226 -----AlaArgArg 228
Db 1064 ATGGACAAACAGTGAACATCTGGCAATTTGACCTGGAAACACTTGGCCAAAGACAGC 1123
Qy 229 ThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluValLysSerThrTrp 248
Db 1124 ACAGACATCATGCTGAGCAAGCAATTTACCGAAGATTTGGCAGAGAGATGCTCAACATGG 1183
Qy 249 LeuCysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsnIleAspGly 268
Db 1184 CTTTGTCACAAAGATTTAAAGATCTTGTGTATTTTCAAGATGAATTAACATTGATGGA 1243
Qy 269 LysGluLeuLeuAsnLeuThrLysGluSerLeuAlaAspAspLeuLysIleGluSerLeu 288
Db 1244 AAAGAACTGTGAAATCTTACAAAGAAAGCTGTGCTGATATTTGAAATTTGAATCTCTA 1303
Qy 289 GlyLeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLysValLysSerLeu 308
Db 1304 GGAAGTGTGTAAGAGCTGAGAAATTTGAAGAGCTCAGCAACAAAGTTAAATCCCTT 1363
Qy 309 SerSerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspPro 328
Db 1364 TCTTCAGGAATTTCTGATGAAATTTAATGTCCTCAATAGAAATCTTATGAAAGATCCG 1423
Qy 329 ValIleAlaSerAspGlyIleProSerIleGluGluAlaMetGluAsnTrpIleSerLys 348
Db 1424 GTCATGCCATCAGATGGCTATTCATATGAAAGAGCAATGAAATTTGGATCAGCAAA 1483
Qy 349 LysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db 1484 AAGAAAGTACAGATCCCATGACAAATCTTGTCTTCCTTCACGGGTAATTACACCAAAAT 1543
Qy 369 ArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGluLys 384
Db 1544 AGGACTCTGAAATGGCCATCAATAGATGGCTGGAGACACCAAAAG 1591

RESULT 7

AKI29983
LOCUS AKI29983 1919 bp mRNA linear PRI 10-SEP-2003
DEFINITION Homo sapiens cDNA FLJ26473 f1s, clone KDN04521.
ACCESSION AKI29983
VERSION AKI29983.1 GI:34526702
KEYWORDS oligo capping; f1s (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Bakayocae; Metazoa; Chordata; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S.,
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Suzuki, Y., Hata, H.,
Nakagawa, K., Mizuno, S., Morinaga, M., Kawamura, M., Sugiyama, T.,
Irie, R., Otsuki, T., Sato, H., Nishikawa, T., Sugiyama, A.,
Kwakami, B., Nagai, K., Isegai, T. and Sugano, S.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1919)
AUTHORS Sugano, S. and Suzuki, Y.
TITLE Direct Submission
JOURNAL Submitted (31-JUL-2003) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Laboratory of Genome Structure, Human Genome
Center, Shitokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan
E-mail: flcdna@img.u-tokyo.ac.jp, Tel:81-3-5449-5286,
Fax:81-3-5449-5416)
COMMENT NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction and 5'-end one pass sequencing: Institute of Medical
Science, University of Tokyo, Laboratory of Genome Structure, Human
Genome Center; 3'-end one pass sequencing: RAB; clone selection for
full insert sequencing: RAB and Helix Research Institute.
FEATURES
source
location/Qualifiers
1..1919
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="KDN04521"
/tissue_type="kidney"
/clone_lib="KDN"
/note="Cloning vector: pME18SFL3"
ORIGIN
Alignment Scores:
Prod. No.: 6.62e-172 Length: 1919
Score: 1978.00 Matches: 382
Percent Similarity: 80.25% Conservative: 0
Best Local Similarity: 80.25% Mismatches: 2
Query Match: 96.63% Indels: 92
DB: Gaps: 1
US-10-077-111-13 (1-384) x AKI29983 (1-1919)
QY 1 MetVallyaleuileuilethrlLeuAlaAspHisGlyAspAspValaAsnCyseValaAlaPhe 20
Db 258 ATGGTGAACCTGATTCACACATTAGCTGATCATGTGATGACATGTCACATGCTGCTTC 317
QY 21 SerPheSerLeuLeuAlaThrCyseSerLeuAspIleThrIleArgLeuIYrSerLeuArg 40
Db 318 TCCTTTTCCTCTGGCTACTTGTCTCTTGGACAAACAAATTCCTCCCTGACTCTTTCGCT 377
QY 41 AspPheThrGluLeuPProHisSerProLeuIlyAspPheIsthrTYrAlaValHisCyse 60
Db 378 GACTTTCCTGAACTGCCACATTTCTCCATTGAAGTTTCATACCTATGCTGCTCCACTGCTGC 437
QY 61 CysePheSerProSerGlyHisIleLeuAlaSerCyseSerThrAspGlyThrThrValIleu 80
Db 438 TGTTTCCTCCCTTCAGGACATATTTTGGCATCGTGTTCACACAGATGGTACACTGTCTCTA 497
QY 81 TrpAsnThrGluAsnGlyGlnMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 498 TGGAAATACCTGAAAATGACAGATGCTGCGAGTGTATGTAACAGCCTAGTGGCAGCCCTGTG 557

QY 101 ArgValCyseGlnPheSerProAspSerThrCyseLeuAlaSerGlyAlaAlaAspGlyThr 120
Db 558 AGGGTTTCCAGATTTTCCCCAGATCTCCAGTGTGGTCATCAGGGGCACTGATGGAACCT 617
QY 121 ValValLeuThrAsnAlaGlnSerTYrIlyLeuTYrArgCyseGlySerValIlyAspGly 140
Db 618 GTGGTTTGTGGAAATGACAGATCATCAAAATTAATAGATGGTATGATTAAGATGGCC 677
QY 141 SerLeuAlaAlaCyseAlaPheSerProAsnGlySerPhePheValThrGlySerSerCyse 160
Db 678 TCCTTGGCCGCGATGTGCATTTCTCTTATATGAAAGCTTTTGTGACTGCTGCTCATGT 737
QY 161 GlyAspLeuThrValTrpAspAspIlyMetArgCyseLeuHisSerGlyIlyAlaHisAsp 180
Db 738 GGTGATTTAAACGTGTGGATGATTAATAATGAGGTGTCTCATATGTAAGTAAACACATGAT 797
QY 181 LeuGlyIlyLeuThrCyseAspAspPheSerSerGlnProValSerAspGlyGluGlnGlyLeu 200
Db 798 CTGGAAATTAACCTGCTGCGATTTTCTTACAGCCAGTTTCTGATGGAGAACAGGCTTT 857
QY 201 GlnPhePheArgLeuAlaSerCyseGlyGlnAspCyseGlnValIlyIleTrpIleValSer 220
Db 858 CAGTTTTCCTGACTGSCATCATGTGTGTCAGATTCGCAAGTCAAAATTTGCATGTCTCT 917
QY 221 PheThrHisIleLeu----- 225
Db 918 TTTAACCATATCTTAGTTTGTGAATTAATTAATAAGTACACTGAGTGAGCATGTGCT 977
QY 225 ----- 225
Db 978 CTTGTTTGGCTTGTGCTTTTTCCTCCATGATGGCAGATCTAGTCTCAGGGTCATGTGAT 1037
QY 225 ----- 225
Db 1038 AAGTCTGTCACTAGTATATGATACTAATACTGAGAAATATACTTCACATTTGACTAGCAC 1097
QY 225 ----- 225
Db 1098 ACCAGGTATGTCAACATTTGTGCTTTTGGACCTAATACCCTTTTACTGCTACTGTGTCA 1157
QY 226 -----AlaArg 228
Db 1158 ATGACAAACAGCTGAACATCTGGCAATTGACCTGGAAACATTTTCCAGACGACAGAC 1217
QY 229 ThrGluHisGlnLeuIlyGlnPheThrGluAspTrpSerGlyGluValIleSerThrTrp 248
Db 1218 ACAAGACATCAGCTGAAGCAATTTTACCAAGATTGCTCAGAGAGAGATGCTTCACATGG 1277
QY 249 LeuCyseAlaGlnAspLeuIlyAspLeuValGlyIlyPheIlyMetAsnAsnIleAspGly 268
Db 1278 CTTGTGACACAAAGATTTAAAGATCTTGTGGTATTTTCAAGATGAATTAACATTGATGGA 1337
QY 269 IlyeGluLeuLeuAsnLeuThrIlyeGluSerLeuAlaAspAspLeuIlyIleGluSerLeu 288
Db 1338 AAAAGAACTGTGAATCTTCAAAAGAAAGCTGTGGCTGAGATTTGAAAATTGATATCTCA 1397
QY 289 GlyLeuAspSerIlyValIleuArgIlyIleGluGluLeuArgThrIlyeValIlyeSerLeu 308
Db 1398 GGACTGGTATGTAAGTCTGAGGAAATTTGAAGACTCAGACCAAGTTAAATCCCTT 1457
QY 309 SerSerGlyIlyeProAspGluPheIleCyseProIleThrArgGluLeuMetIlyAspPro 328
Db 1458 TCTTACGGAATTCCTATGATTAATTAATGTCATTAACATTAAGAACTTATGAAAGATCCG 1517
QY 329 ValIleAlaSerAspGlyTYrSerTYrGlyIlyGluAlaMetGluAsnTrpIleSerIlye 348
Db 1518 GTCATGCGATCAGATGCTATTCATATGAAAAGAAAGCAATGGAATTTGGATCAGACAAA 1577
QY 349 IlyeIlyAspArgThrSerProMetThrAsnLeuValLeuProSerAlaValLeuThrProAsn 368
Db 1578 AAGAAACGTACAAAGTCCCATGACAAATCTTGTCTTTCAGCGGTACTTACACCAAAAT 1637

OY 369 ArgThrLeuIysMetAlaIleAsnArgTrrpLeuGluThrHisGlnLys 384
Db 1638 AGGACTCTGAATAATGGCCATCATATAGATGGCTGGAGACACCAAAAG 1685
RESULT 8
CO729621
LOCUS CO729621 1254 bp DNA linear PAT 03-FEB-2004
DEFINITION Sequence 15555 from Patent WO2068579.
ACCESSION CO729621
VERSION CO729621.1 GI:42301208
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
REFERENCE
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.
TITLE Kites, such as nucleic acid arrays, comprising a majority of
humanexons or transcripts, for detecting expression and other uses
thereof
JOURNAL Patent: WO 02068579-A 15555 06-SEP-2002;
PE Corporation (NY) (US)
FEATURES
source Location/Qualifiers
1..1254
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores:
Pred. No.: 6,13e-158 Length: 1254
Score: 1823.50 Matches: 357
Percent Similarity: 85.95% Conservative: 4
Best Local Similarity: 85.00% Mismatches: 20
Query Match: 89.08% Indels: 39
DB: Gaps: 4
US-10-077-111-13 (1-384) x CO729621 (1-1254)
OY 1 MetValIysLeuIleHisThrLeuAlaAspHisGlyAspAspValaAsnCyseValaAspHe 20
Db 1 ATGTGTGAACATGATTCACATATAGCTGATCATGTGTGACATGTCATCTGTCCTTC 60
OY 21 SerPheSerLeuLeuAlaThrCyseSerLeuAspLysThrIleArgLeuYrSerLeuArg 40
Db 61 TCCCTTTCCTCTGCTGCTACTGCTCTCTGACAAACAAATTCCTGCTGCTGCTGCTGCT 120
OY 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrIleAlaIleCyseCys 60
Db 121 GACTTACTGAACTGCCACATTCCTCATTTGAAATTTCTATACCTATGCTGCTGCTGCTG 180
OY 61 CysePheSerProSerGlyHisIleLeuAlaSerCyseThrAspGlyThrThValleu 80
Db 181 TGTTCCTCCCTTACAGACATATTTTGGCATCGTTTCACACAGATGTATCCACCTGCTCTA 240
OY 81 TrpAsnThrGluAsnGlyIleMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 241 TGGAACTGTGAAATGAGACATCTGCGACATGAGAACGCTTAGGGGAGCCCTGTG 300
OY 101 ArgValCyseGlnPheSerProAspSerThrCyseLeuAlaSerGlyAlaAlaAspGlyThr 120
Db 301 AGGGTTTGGCAGTTTCCCGACACTCCACGCTTTGGCATAGGGGAGCGTGAATGAACT 360
OY 121 ValValLeuTrpAsnAlaGlnSerTrpLysLeuTrpArgCyseGlySerValLysAspGly 140
Db 361 GTGGTTTGTGGATGACACTCTATACAAATTAATATATGATGTGTATTAAGATGGC 420
OY 141 SerLeuAlaAlaCyseAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160
Db 421 TCCCTTGGGCGCATGTCATTTCTCTATATGAAAGCTTCCTTGTACCTGCTCTCAAGT 480
OY 161 GlyAspLeuThrValTrpAspAspLysMetArgCyseLeuHisSerGlyLysAlaHisAsp 180

Db 481 GGTGATTTAAACAGTGTGGGATGATTAATAATGAGGTGTCTGATAGTAAAGACATGAT 540
OY 181 LeuGlyIleThrCyseAspPheSerSerGlnProValSer----- 194
Db 541 CTTGGAAATTAACCTGCTCGATTTTCTTCACACAGATTTCTGTTTGAATTAATAAT 600
OY 195 -----AspGly 196
Db 601 AAAAGTACACTGATGGGACACTGTGCTCTGTTCTGGCTTGTCTTTTCCCATATGGG 660
OY 197 GluGlnGlyLeuGlnPhePhe-----ArgLeu 205
Db 661 CAGATGCTATGCTCAGGGTATGTCACAACTTGCTTTTGCACCTAATACCTTTTACTT 720
OY 206 AlSerCyseGlyGlnAspCyseGlnValLysIleTrpIleValSerPhe---ThrIle 224
Db 721 GCTACTGTTCATATGACAAACAGTACATCTGCAATTGACCTCGAAACACTTTC 780
OY 225 LeuAlaArgArgThrGluHisGlnLeuLysGlnPheThrGluAspTrpSerGluGluVal 244
Db 781 CAAGCAAGGAGCACAAGAACATCAGCTGAGCAATTAACGAAGATTGTCAGAGAGAGAT 840
OY 245 ValSerThrTrpLeuCyseAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetAsn 264
Db 841 GTCTCAACATGCGCTTTGTGACAAAGATTTAAAGATCTTGTGTATTTTCAAGATGAAT 900
OY 265 AsnIleAspGlyLysGlyLeuLeuAsnLeuThrGluSerLeuAlaAspAspLeuLys 284
Db 901 AACATTGATGAAAGAACTGTGAACTTTCACAAAGAAAGCTGCTGATGATTTA 960
OY 285 IleGluSerLeuGlyLeuArgSerLysValLeuArgLysIleGluGluLeuArgThrLys 304
Db 961 ATT-----GGACTGGCTAGTAAAGTCTGAGAAATTTGAAGAGCTCAGAGACCAAG 1011
OY 305 ValLysSerLeuSerSerGlyIleProAspGluPheIleCyseProIleThrArgGluLeu 324
Db 1012 GTTAAATCCCTTTCTTCAGAAATTCCTGATGAATTTATATGCTCCAAATACCTAGGAAT 1071
OY 325 MetLysAspProValIleAlaSerAspGlyLysSerTrpGlyLysGluAlaMetGluAsn 344
Db 1072 ATGAAGATCCGGTATCGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131
OY 345 TrpIleSerLysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaVal 364
Db 1132 TGGATCGCAAAAAGAACTACATCAAGTCCATGACAAATTTGTTCTCTTACGGGCTA 1191
OY 365 LeuThrProAsnArgThrLeuLysMetAlaIleAsnArgTrpLeuGluThrHisGlnLys 384
Db 1192 CTTAACCAATATGACTCTGAAATGGCCATCATATGATGAGCTGGAGACACCAAAAG 1251
RESULT 9
BC061948 2130 bp mRNA linear ROD 19-NOV-2003
LOCUS BC061948
DEFINITION Mus musculus RIKEN cDNA 2610014F08 gene, mRNA (cDNA clone MGC:65547
IMAGE:6312489), complete cds.
ACCESSION BC061948
VERSION BC061948.1 GI:38303926
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buote, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schneitz, T.E., Brownstein, M.J., Uebachs, T.B., Tosh, Yuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N., Yoshinuki, S.,
Abramson, R.D., Muliyil, S.J., Boseak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Sanchez, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchéz, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257
12477932
2 (bases 1 to 2130)
Strausberg, R.
Direct Submission
Submitted (10-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.ncl.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Susan L. Sullivan, PhD.
CDNA Library Preparation: Reggen, Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nsl.nih.gov
Ahter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stanki, P., Thomas, P.J., Touchman, J.W., Tsaur, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
Series: IRAC Plate: 127 Row: 1 Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES
source
Location/Qualifiers
1. 2130
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:65547 IMAGE:6312489"
/tissue_type="Olfactory epithelium, neonatal mouse,
C57Bl/6"
/clone_1ib="NIH_MGC_129"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 2130
/gene="2610014F08R1k"
/note="synonym: 1700048E19R1k"
/db_xref="LocusID:72137"
/db_xref="MGI:1919387"
208. 1632
/gene="2610014F08R1k"
/codon_start=1
/product="2610014F08R1k protein"
/protein_id="AAH61948.1"
/db_xref="GI:38303927"
/db_xref="LocusID:72137"
/db_xref="MGI:1919387"
/translation="MRLITHTLADHDGDDVCCAFSAALATCSLIDKTRIRYSIDPAE
LPYSPLFHTYAVHCCFSPSGVHVASCSIDGTIVMSSHSHTLVLEPQGSFVAV
CPSPSDAVYLASAGADGSIAMNAQTKLVRCGSYVDSIVACAFSPDGLFVTGSSG

CDLTVWDBMRCLHSEKADLGTCCSRSSQPLSGEGIQSYQLASCCDCEIKMAY
TTTRVGLFELKXKSTLSGCACAVLACASHDDKMLASGVDSVLIHGIGPQSVHTL
TQTRVLTTCAPAPNTLLATGSMKXTVNIWQFDETPQGSMDPLKPIETWSEED
VAVWRAQGLDGLVGIFRANIDGKELHLITRESLAGDKISLGRSKYLSIELIR
AKDMSAGSIPDEPFCPTRELMKQPVASDGYSYREAMESEIHKKKTSPTNLAL
BELVLTNPNTLMCAINRWLETHK"
217. 1131
/gene="2610014F08R1k"
/note="WD40; Region: WD40 domain, found in a number of
eukaryotic proteins that cover a wide variety of functions
including adaptor/regulatory modules in signal
transduction, pre-mRNA processing and cytoskeleton
assembly"
/db_xref="CCD:cd00200"
1192. 1383
/gene="2610014F08R1k"
/note="SAM; Region: Sterile alpha motif."
/db_xref="CCD:cd00166"
1408. 1629
/gene="2610014F08R1k"
/note="U-box; Region: U-box domain. This domain is related
to the Ring finger pfam00097 but lacks the zinc binding
residues"
/db_xref="CCD:pfam04564"

ORIGIN
Alignment Scores:
Pred. No.: 1,236-138 Length: 2130
Score: 1617.00 Matches: 312
Percent Similarity: 72.21% Conservative: 31
Best Local Similarity: 65.68% Mismatches: 40
Query Match: 78.99% Indels: 92
DB: 10 Gaps: 3

US-10-077-111-13 (1-384) x BC061948 (1-2130)
QY 1 MetValIleuLeuIleIleThrLeuAlaAaPHISGLYAspAspValaIaenCyCyAlaPhe 20
Db 208 ATGGTAGGCTGATTCACACGCTGCTGATACCGGAGTAGAGTCACTGCTGCGCTTC 267
QY 21 SerPheSerLeuLeuAlaThrCySerSerLeuAlaPylValThrIleArgLeuYrSerLeuArg 40
Db 268 TCGGCTGCGCTTCTGCGCACCTGCTCTTGACACAGACATCGGCTGATCTCCCTAAGT 327
QY 41 AaPheThrGluIleuProHISerProLeuIlePheIleThrTyraIaValHisCyCyAs 60
Db 328 GACTTTCGTAAGTCCGCTGATCTCCCGCTGAAGTTCCACACTATCTCTCACTGCTCG 387
QY 61 CyAPheSerProSerGlyHisIleLeuAlaSerCySerThrAspGlyThrValIleu 80
Db 388 TGTTCCTCACCTCAGACACAGTTTAAAGATCGTGTCCAGACAGCGGAGCTGCTG 447
QY 81 TrpAsnThrGluAsnGlyIleMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 448 TGGAGCTCCACACGCGGACACACCTGACCGGTGTGAGACACGCGGTGCGACGCTGTG 507
QY 101 ArgValCyGlnIlePheSerProAspSerThrCySerLeuAlaSerGlyValIaIaAspGlyThr 120
Db 508 CGGCTGTGCTGCTTCCCAAGACTGCTGCTACTGCTGAGGGCTGCGCATGATTC 567
QY 121 ValIaIleuTrpAsnIaIaIleSerTyryIleuTyraIaGlySerValIaIaAspGly 140
Db 568 ATGCTTGTGGATGACACAGACATATCAAACTATATAGATGTGTGATCAAGATAC 627
QY 141 SerLeuAlaIaCyAlaIaPheSerProAsnGlySerPhePheValThrGlySerSerCy 160
Db 628 TCATTGTGTGCGCTGCTTCTCCCAAGAGGCTCTTGTATCACTGCTCTCGGCG 687
QY 161 GlYAspLeuThrValIaIaPheAspIleuIleMetCysGlySerValIaIaIaAsp 180
Db 688 GGGAGACTTGACACTGTGGATGACAGATGAGGTGTCTACACAGCAGAGGCGCACAT 747
QY 181 LeuGlyIleThrCyAspAspPheSerSerGlnProValSerAspGlyIleGluGlnIleu 200


```
Db      748 CTCGGAGATCACCTCTGAGCGCTTTCCACAGCGCTCTCTGGCGGAGAA---GGCTTC 804
Qy      201 GlnhphearglaualaserCygllyglnapcygslvalyslletrp11levalaser 220
      805 CAGCTTACAGGTGGCGTCATGTGTCAAGACTGTGAATCAAACTCTGGGCTGTACT 864
Qy      221 PheThr----- 222
      865 ATTACCGCGTCTTAGCGTTGAATTAATAATAAAGCACACTAAGTGGCAGTCGCC 924
Qy      222 ----- 222
      925 CCTGTCTGCGCTGTGCTTTTTCACATGATGAAAGATCTTGCACTGGGCTCAGTGAT 984
Qy      223 -----His1leuAlaArg 228
      985 AAATGTGTCATCATACATGATGTGGCCCTCAGAGTGTGTACACACGCTGACTCAGCAT 1044
Qy      229 ThrGlnHis----- 231
      1045 ACCAGGATGTGTAAGACTTGTGCGCTTGACCCCAACACTCTTACTTGTACTGTTC 1104
Qy      231 ----- 231
      1105 ATGACAAGACAGTGAACATTTGGCAGTTTGACCTGAGAAACACCTTGCCAGAGAGATG 1164
Qy      232 -----GlnleuLyglnPheThrGlnAspTrpSerGlnGluValSerThrTrpLeu 249
      1165 AACGACCCGCTGAAACATTTCACTGAGAAAGTGTGTCAGAGAGAGATGTCCTCGTGTGCTT 1224
Qy      250 CysAlaGlnAspLeuLygAspLeuValGlyTrpPheLygMetAsnAsn1leAspGly 269
      1225 CGTGTCTAAGCGCTTGGAGACCTGTCGCTGATTTTCAGGGCAACAAATGATGAGGAAA 1284
Qy      270 GlnleuLeuAsnLeuThrLygGlnSerLeuAlaAspAspLeuLyg1leGlnSerLeuGly 289
      1285 GAACATATGATCATCAACAAAGAAAGTGTGCTGTGATTTGAAATGATGATCTCTAGG 1344
Qy      290 LeuArgSerLygValLeuArgLyg1leGlnGlnLeuArgThrLygValLygSerLeuSer 309
      1345 CTGGCGACGAAAGCTCTGAGAGATTTGAAAGCTCAGGGCGCAAGATGATGATTCCTCTCT 1404
Qy      310 SerGlyTrpProAspGlnPhe1leCysPro1leThrArgGlnLeuMetLygAspProVal 329
      1405 TCCGGAATCCCTGACGAGTATCTGCCCCAATACAGAGAACTCATGAAGACCCGCTC 1464
Qy      330 1leAlaSerAspGlyTrpSerLygGlnLygGlnAlaMetGlnAsnTrp1leSerLygLy 349
      1465 ATGCATCATAGTGTACTCTCTACGAGAGAGAGAAAGATGGAAGCTGATCCACAGAAG 1524
Qy      350 LygArgThrSerProMetThrAsnLeuValLeuProSerAlaVal1leuThrProAsnArg 369
      1525 AAGGCTACGAGCCCATGACAAATTTTGCTCTCCCTTCACTGATGACCCCAACAGG 1584
Qy      370 ThrleuLygMetAla1leAsnArgTrpLeuGlnThrHisGlnLyg 384
      1585 ACACCTGAAGATGGCCATCAACCGATGCTGAGAGACGACAGAGAG 1629

RESULT 10
LOCUS      AJ719461          3818 bp      mRNA      linear      VRT 30-SEP-2004
DEFINITION Gallus gallus mRNA for hypothetical protein, clone 2121.
ACCESSION  AJ719461
VERSION     AJ719461.1 GI:53127473
KEYWORDS   ORF1.
SOURCE      Gallus gallus (chicken)
ORGANISM   Gallus gallus
            Bukarjota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
            Phasianinae; Gallus.
REFERENCE   1
AUTHORS    Caldwell,R.B., Kierzek,A.M., Arakawa,H., Bezubov,Y., Zaim,J.,
            Fiedler,P., Kuter,S., Blagoderaki,A., Kostovska,D., Koter,M.,
            Plachy,J., Carnici,P., Hayashizaki,Y. and Buerstedde,J.M.
            Full-length cDNAs from bursa lymphocytes to facilitate gene
            function analysis
            Unpublished
            2 (bases 1 to 3818)
REFERENCE   2
AUTHORS    Caldwell,R.B.
            Direct Submission
            Submitted (20-MAY-2004) Caldwell R.B., GSF - Forschungszentrum,
            Institut fuer Molekulare Strahlenbiologie, Ingolstaedter Landstr.
            1, D-85764 Neuherberg, GERMANY
            Location/Qualifiers
            1..3818
            /organism="Gallus gallus"
            /mol_type="mRNA"
            /beta_in="CB"
            /db_xref="taxon:9031"
            /clone="2121"
            /cell_type="lymphocyte"
            /tissue_type="bursa"
            /clone_1ib="riken1"
            /dev_stage="2 weeks old"
            1..3818
            /locus_tag="RCJMB04_2121"
            <1..341
            /locus_tag="RCJMB04_2121"
            342..1469
            /locus_tag="RCJMB04_2121"
            /note="ORF1"
            /codon_start=1
            /product="hypothetical protein"
            /protein_id="CA031120.1"
            /db_xref="GI:53127474"
            /translation="MATLIRHTLDHSDDVNCCAFSSSCLATGSLDKTIRIYSLDPTF
            LPSPLKGTVAHCCCFSPSGHTLASCTSDGATIIWDSDEMLAVLEQPTSPVAV
            CRSPSPSTLYVSGADGSLVLMVHSMKPYRSGNDSGLVACA FSGGPFYTGSSC
            GDLTMDDKRCCENKRAHDLGVTCDSISHPYSDGSHAGCFOMASCGDGNKIKWF
            ILPADFLGCTLRKCTLSGHSRAVLTCASVTDQOMVSGSVDCVITVYENNTLIHT
            LSGHRTVTTTCARAPCSLPLATSGSMKTHIWKLDNKQPCAGNTIENDSKIRIDENMS
            EDVSAWLCAGQFAELVGLFRANNIDGKELVNLTRSLIHE"
            1470..>3818
            /locus_tag="RCJMB04_2121"

ORIGIN
Alignment Scores:
Pred. No.:      3,18e-124      Length:      3818
Score:          1464.00      Matches:      281
Percent Similarity: 67.30%      Conservative: 40
Best Local Similarity: 58.91%      Mismatches: 62
Query Match:    71.52%      Indels:      94
DB:              5      Gaps:      3

US-10-077-111-13 (1-384) x AJ719461 (1-3818)
Qy      1 MetValLygLeu1leHisThrLeuAlaAspHisGlyAspAspValAsnCygSalaPhe 20
      342 ATGGCGACATTAATTCACACTTTAAACAGATCAAGTATGATGCAACTGTGCGCTTC 401
Qy      21 SerPheSerLeuAsnAlaThrCysSerLeuAspLygThr1leArgLeuTrpSerLeuArg 40
      402 TCATCATCGTGTGGTGTACTTGTCTTGTGACAAACAATTTGGTATTTACTCTTTAAAT 461
Qy      41 AspPheThrGlnLeuProHisSerProLeuLygPheHisThrTrpAlaValHisCygS 60
      462 GACTTCACATAGCTCCCATACTCACCTTTGAAAGTATCATATGCTGTGCCACTGTGC 521
Qy      61 CysPheSerProSerGlyHis1leLeuAlaSerCysSerThrAspGlyThrTrpValLeu 80
      522 TGCTTCTCTCATCAGACACACTTATGCTTCGTTGTTCAACACAGCGGTGCTTACCATAT 581
Qy      81 TrpAsnThrGlnAsnGlyGlnMetLeuAlaValMetGlnGlnProSerGlySerProVal 100
      582 TGGGACACTAGCGATGTGTGATGCTGTGGCAGTGTGAGCAGCCACTGGCAGTCTCTGC 641
```

```

Qy 101 ArgValCysGlnPheSerProaspSerThrCysLeuAlaSerGlyAlaAlaaspGlyThr 120
Db 642 AGAGTGTGCGCGCTTTTCTCCGAGTCCATTACTGGTGTCAAGGTGACGTGATGAGAT 701
Qy 121 ValValLeuThrPheAlaGlnSerTyrLysLeuTyrArgCysGlySerValLysaspGly 140
Db 702 GTGTCTTCTTGGAAAGTGCACCTCAATGAAATTTCTACGATCTGGGAATGTTAAAGATGT 761
Qy 141 SerLeuAlaAlaCysAlaPheSerProaspGlySerPhePheValThrGlySerSerCys 160
Db 762 TCTTTGGTGGCGCTGTGCTTTTCTCCAGGTGGAATTTCTTTGTACTGATCTCATGCT 821
Qy 161 GlyAspLeuThrValTTPaaspApLyMecArgCysLeuHisSerGlyLysValAsp 180
Db 822 GGTGATCTGACTGTTTGGATGATTAATGATGATGCTCTGTATATGAAAAAGCATGAT 881
Qy 181 LeuGlyLysThrCysCysaspPheSerSerGlnProValSeraspGlyGlnGlnGlyLeu 200
Db 882 CTGGCGCTTACCTGCTGTGATATTCTTTCACATCCATCTTGATGGGAAACATGCAATCC 941
Qy 201 GlnPhePheArgLeuAlaSerCysGlyGlnaspCysGlnValLysIleTPIIeValSer 220
Db 942 GGAATGCTTCCAGATGCGATCTTGTGTCAAGATTAATGAATCAAAAGTCTGTTATTG 1001
Qy 221 PheThrHisIleLeuAlaArgArgThrGlu----- 230
Db 1002 TTTCAGACACTTTAGGTGGAATTGAGATATTAATGACATTGAGTGCACATTCTGCC 1061
Qy 230 ----- 230
Db 1062 CCAGTTCTGACTGTGCAATTTCTTATGATGGCGAGATGTGGTTTCAGGGCTGTGAC 1121
Qy 231 -----HisGlnLeuLysGlnPhe 236
Db 1122 AAGTGTCTAATATATGAGACTAATATCGGCATATCTCTTCACTTGTCTCAGCAT 1181
Qy 237 Thr----- 237
Db 1182 ACTAGATATGTCACAACTGTGCTTGTCTCATGTAGTGTCTTCTTGCCAGCGGTTC 1241
Qy 237 ----- 237
Db 1242 ATGATATAAACTGTGCACATATGGAACCTGACATTAAGAACCTGTGACGAACACT 1301
Qy 238 -----GluaspTrpSerGlnGluValLysThr 247
Db 1302 ATGAAAAATGACTCTTAAGATTAGAACTGTGAGAACTGTGAGAGATGATGTTTCAGCC 1361
Qy 248 TrpLeuCysAlaGlnaspLeuLysaspLeuValGlyLysPheLysMetAsnAsnLysasp 267
Db 1362 TGGCTTTGTGCACAAGGCTTGCAGAACTTGTGGGCTTTTCAAAAGCAATTAATCATGAT 1421
Qy 268 GlyLysGlnLeuLeuAsnLeuThrLysGlnSerLeuAlaaspLeuLysIleGlnLys 287
Db 1422 GGCAGAAAGACTCTGGAATCTTAACAAGAAAGCTGATTCATGATTAATAAAGT---AGT 1478
Qy 288 LeuGlyLysLeuArgSerLysValLysArgLysIleGlnGlnLeuAsnThrLysValLysSer 307
Db 1479 CTCTGGCTGGCGCAATTAATCTCTCAAGAAATGCAAGCTGAGATGAAATGCTCTCT 1538
Qy 308 LeuSerSerGlyLysProaspGlnPheLysCysProLysThrArgGlnLeuMetLysasp 327
Db 1539 GTTCTGCTGTGCTGCTGCTGATGATTTCTTATGCCCTTAACAAGGAGCTCATGAAGAT 1598
Qy 328 ProValLysLeaSeraspGlyLysSerTyrGlnLysGlnAlaMetGlnAsnTPILeser 347
Db 1599 CTTGTCAATTCAGACAGATGGGTATTCTATGAAAAAGAGCAATGAGAAATTTGATCAGC 1658
Qy 348 LysLysLysArgThrSerProMetThrAsnLeuValLeuProSerAlaValLysLeuThrPro 367
Db 1659 AATTAACGACGATCTAGTCCCATGACAAATCTTCTCTCCAGCTTGTGCTTAACCC 1718

```

```

Qy 368 AsnArgThrLeuLysMetAlaIleAsnArgTTPLeuGlnThrHisGlnLys 384
Db 1719 AACAGACACTGAAATGCTATCAGCCCTGCTGAGACATCAGCAAAA 1769

RESULT 11
LOCUS
BC050792
DEFINITION
Mus musculus RIKEN cDNA 2610014F08 gene, mRNA (cDNA clone
IMAGE:6742025) , partial cds.
ACCESSION
BC050792
VERSION
BC050792.1 GI:30048136
KEYWORDS
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
1 (bases 1 to 1547)
Serauberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marsina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Uedtin,T.B., Tothylaki,S.,
Carninci,P., Brownstein,M.J., Uedtin,T.B., Tothylaki,S.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
Mckernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hultk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,J.C., Snevchenko,Y.,
Bouffard,G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Sklarska,U., Smallus,D.B.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26) , 16899-16903 (2002)
2 (bases 1 to 1547)
Serauberg,R.
Direct Submission
Submitted (09-APR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Dr. Jonathan Kuo, NIMH
cDNA Library Preparation: Michael Brownstein / Ted Uedtin
laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Boedeltje, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hailao, Martin Krzywinski, Reia Kutche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice Mcleavy, Steven
Ness, Pawan Parbhu, Anna-Lisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smallus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tasi, Natasja van den Bosch, Jill Verdy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 46 Row: a Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
Location/Qualifiers

```

source
1. .1547
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6742025"
/tissue_type="Testicle, mouse"
/clone_id="N1H_MGC_169"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
<1. .1547
/gene="2610014F08Rik"
/note="synonym: 1700048E19Rik"
/db_xref="LocusID:72137"
/db_xref="WGI:1919387"
<1. .1442
/gene="2610014F08Rik"
/codon_start=3
/product="2610014F08Rik protein"
/protein_id="AAH50792.1"
/db_xref="GI:30048137"
/db_xref="LocusID:72137"
/db_xref="WGI:1919387"
/translation="PSSGSPSPGQCRGLDTRRQCVTSTPEPRRGAARRRDS
GCRGSCILCVNWRILHTLADHDVSCAFSALATSLDKTILYSLSPAEIP
YSPLEHTYAVHCCFSPSGHVLASCTDGTVMSSHGHITLVLEPGSGPRVC
FSPSAYLAGAAGSIALNMAOTYKLYRGSVSDSLVCAFPDGLFVTSGGD
LYWDDMRCLHSEKADHLCITCSRSSQSGSGEGQVQLASCGGDCIKIMAVT
TRYVPELTKSTLSHGCAPYLACAFSHDGRMLASGVDSVLIHGIQPSVHTLTQ
HTRVTTCAFPNTLLATGSMKTANIWQFDLETPOGSMNDPLKFTMESEDS
VMLRAQGLIEDLVGIFRANNIDGKELHLHTKESLAGDUIKISLGRSKVRSIEBLRAK
MDSISGIPDEFICPTIRELMKDPVIASG"

Alignment Scores:
Pred. No.: 2.25e-116 Length: 1547
Score: 1373.00 Matches: 272
Percent Similarity: 67.18% Conservative: 33
Best Local Similarity: 59.91% Mismatches: 45
Query Match: 67.07% Indels: 104
Gaps: 4

US-10-077-111-13 (1-384) x BC050792 (1-1547)

QY 1 MetValysLeuIleHsThrLeuAlaAspHisGlyAspAspValaAnCyCysAlaPhe 20
DB 171 ATGATGAGGCTGATTCACGCTGCTGATACGCTGATACGCTGCTGCTGCTTC 230
QY 21 SerPheSerLeuLeuAlaThrCysSerLeuAspThrIleArgLeuYrSerLeuArg 40
DB 231 TCGGCTGCTCCTCCGCGCACCTGCTCTTGACAAAGACATCCGCTGCTCTCACTAGT 290
QY 41 AspPheThrGluLeuProHisSerProLeuAspPheHisThrYrAlaValHisCysCys 60
DB 291 GACTTCTGTAACGCGTACTCCCGCTGAACTTCAACACTATGCTGCTCACTGCTGC 350
QY 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
DB 351 TGTTCACCTCAGACAGGTTTGTAGCATGCTGCAAGACGAGGACACGAGTCTG 410
QY 81 TrpAsnThrGluAsnGlyIleMetLeuAlaValMetGluGlnProSerGlySerProVal 100
DB 411 TGAAGCTCGACAGCAGACACACCTTACCTGTGAGACAGCCGGGTGGAGCCCTGTG 470
QY 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
DB 471 CGGCTCTGTTGCTTTCCCGAGACTGCTTACCTACGCTACAGGCGCTGCCGATGATCC 530
QY 121 ValValLeuThrAspAsnAlaGlnSerThrYrLeuLeuYrArgCysGlySerValValAspGly 140
DB 531 ATTGCTTTGGATGACAGACATACAACTATATAGGTGTGTATGTCAAGATGAC 590
QY 141 SerLeuAlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCys 160

DB 591 TCATTGTGCTGCTGCTGCTTTCTCCGATGAGAGGCTCTTTGTCACTGCTCTCGGC 650
QY 161 GlyAspLeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAsp 180
DB 651 GGGGACTTGAAGTGGATGACAGATAGAGTGTCTACACAGCAGAAAGGCGACGAT 710
QY 181 LeuGlyIleThrCysCysAspPheSerSerGlnProValSerAspGlyGlnGlyLeu 200
DB 711 CTGGGATTCACCTGCTGCTGCTTTCTCTCACTGCTCTCTCTGCGGAGAA---GGCTC 767
QY 201 GlnPheAspLeuAlaSerCysGlyGlnAspCysGlnValIleThrPheValSer 220
DB 768 CAGCTTACAGATTGGCGCTGATGTGTCAAGACTGTAATCAAACTGCGCTGTACT 827
QY 221 PheThr----- 222
DB 828 ATTACCGCTGTAGCTTGAATTAATAATATAAGACACTAAGTGGCACTGCGCC 887
QY 222 ----- 222
DB 888 CTGTTCTGCTGCTGCTTTTTCATGATGAAAGATGCTTGCATCGGGGTACGTGAT 947
QY 223 -----HisIleLeuAlaArgArg 228
DB 948 AAATCTGTCATCATACATGATGATCGGCTCAGAGTGTCTACACAGCTGACTCAGAT 1007
QY 229 ThrGluHis----- 231
DB 1008 ACCAGGATGTGTACGACTGTGCTTGTGACCCCAACACTCTTACTGCTACTGTTCA 1067
QY 231 ----- 231
DB 1068 ATGACAAAGACAGTGAACATTTGGCAGTTGACCTGGAACACCTTCCAGAGACATG 1127
QY 232 -----GlnLeuLysGlnPheThrGluAspTrpSerGluGluValIleSerThrTrpLeu 249
DB 1128 AAGACCCGCTGAACATTTCACTGAAGATGTGTGAGAGGATGTCTCGTGTGCTT 1187
QY 250 CysAlaGlnAspLeuLysAspLeuValGlyIlePheLysMetCysAsnIleAspGlyLys 269
DB 1188 CGTCTCAAGCTTGGAGACCTCGCTGATTTTTCAGGCGCAACAACTGATGAGGAA 1247
QY 270 GlnLeuLeuAsnLeuThrLysGlnSerLeuAlaAspAspLeuLysIleGlnSerLeuGly 289
DB 1248 GAATCTATGATCTCAACAAAGAAAGCTGCTGCTGATTTGAATTCGATCTTACAGG 1307
QY 290 LeuArgSerLysValIleuArgLysIleGluGlnLeuArgThrLysValLysSerLeuSer 309
DB 1308 CTGCGCAGCAAACTCTGAGGAGATTTGAAGACTCAGGCGCAAGATGATCTCTCT 1367
QY 310 SerGlyIleProAspGluPheIleCysProIleThrArgGluLeuMetLysAspProVal 329
DB 1368 TCCGGAATCCCTGACGATTCATCTGCCCAATACAGAGAACTCATGAGAGACCCGCT 1427
QY 330 IleAlaSerAspGlyLysSerThrGlyLysGluAlaMetGlu----- 343
DB 1428 ATGCAATCAGGTTACTAGCAGGTCAAAAGAAATCTCTAATAATTAATGATAA 1487
QY 344 -----AsnTrpIleSerLysLysLeuArg 351
DB 1488 GAATTTGTGATCCACAGAAATTAAGAAAGAAAAA 1529

RESULT 12
AX077675 1291 bp DNA linear
LOCUS AX077675
DEFINITION Sequence 130 from Patent WO0105970.
ACCESSION AX077675
VERSION AX077675.1 GI:13122057
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

PAT 22-FEB-2005

REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
1. 1291
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/note="Incyte ID No: 4546403CB1"

ORIGIN

Alignment Scores:
Pired. No.: 5.29e-103 Length: 1291
Score: 1226.50 Matches: 246
Percent Similarity: 71.97% Conservative: 3
Best Local Similarity: 71.10% Mismatches: 4
Query Match: 59.92% Indels: 93
DB: Gaps: 2

US-10-077-111-13 (1-384) x AK077675 (1-1291)

QY 131 LeuTyrcGyagGlySerValLyAspGlySerLeuAlaCysAlaPheSerProAsn 150
1 CTGAGAGGAAATCGGCTCG---AGAGATGCGCTCTCGGCGGCACTGATTTCTCCAAAT 57
151 GlySerPhePheValThrGlySerSerCysGlyAspLeuThrValITPAspAspLysMet 170
58 GGAAGCTTTCTTTCGACTGCGCATTCATGTGTCATTTAAACGTGTGGATGATCAAAATG 117
QY 171 ArgCysLeuHisSerGlyLysAlaHisAspLeuGlyLeuThrCysCysAspPheSer 130
118 AGGTGCTCGATGATGAAAAAGCAGATCTTGGAATTAACCTGCGCATTTTCTTCA 177
191 GlnProValSerAspGlyGlnGlyLeuGlnPhePheArgLeuAlaSerCysGlyGln 210
178 CAGCCAGTTCTCTGATGAGAAACAGGCTTTCAGTTTTCGACTGCGCATTCATGTGCTGAC 237
QY 211 AspCysGlnValLysIleTPIleValSerPheThrHisIleLeu----- 225
238 GATTGCCAAGTCAAAATTTGATGTTCTTTTACCATATCTTAAGTTTGAATTAAA 297
QY 225 ----- 225
298 TATAAAGTACACTGAGTGGGCACTGTGCTCTGCTGCTGCTGCTTTTCCCATGAT 357
QY 225 ----- 225
358 GGGCAGATGCTAGTCTCAGGGTCACTGAGTAAGTCTGTCAATGATATGATCTAAATCT 417
QY 225 ----- 225
418 GAGAATATCTTCAACATTTGACTCAGACACACAGATATGACAACTGTGCTTTGCA 477
QY 225 ----- 225
478 CTAATACCTTTTACTGCTACTGTTCAATGAGCAAAAACAGTAACATCTGGCAATTT 537
QY 226 -----AlaArgArgThrGlnHisGlnLeuLysGlnPheThrGln 238
538 GACCTGAAAAACCTTTGCCAAGAGAGACAGAAACATCAGCTGAAAGCAATTTTACCGAA 597
QY 239 AspTyrSerGlnGlnValIleSerThrTyrLeuCysAlaGlnAspLeuLysAspLeuVal 258
598 GATTGGTCAGAGGAGATGTCTCAACATGCTTGTGCAACAAGTTTAAAAAGATCTTGT 657
QY 259 GlyIlePheLysMetAsnAsnIleAspGlyLysGlnLeuLeuAsnLeuThrLysGlnSer 278
658 GGATTTTCAAGATGATTAACATTTGATGAGAAAAGAACTGTTGATCTTACAAAAGAAAT 717

QY 279 LeuAlaAspAspLeuLysIleGlnSerLeuGlyLeuArgSerLysValLeuArgLysIle 298
718 CTGGCTGATGATTTGAAATTAATCTAGAGACTCGGTAAGTAAGTCTGAGAAAAAT 777
QY 299 GlnGlnLeuArgThrLysValLysSerLeuSerGlyLysIleProAspGlnPheLys 318
778 GAAGAGCTCAGACCAAGGTAAATCCCTTTTCAAGAAATTCGATGAAATTTATATAT 837
QY 319 ProIleThrArgGlnLeuMetCysAspProValIleAlaSerAspGlyTyrSerTyrGln 338
838 CCAATTAACTAAGAGAACTTTATGAAGATCCGCGCATCGCATGACATGAGCATTAATGA 897
QY 339 LysGlnAlaMetGlnAsnTPIleSerLysLysValArgThrSerProMetThrAsnLeu 358
898 AAGGAAGCAATGGAATAATTGATCAAGCAAAAAGAAACGTAACAGTCCCATGCAAAATCTT 957
QY 359 ValLeuProSerAlaValLeuThrProAsnArgTyrIleLysMetAlaIleAsnArgTyr 378
958 GTTCTTCTTCAGCGGCTACTTACACCAATAGGACTGTAATAATGCCATCAATATGATG 1017
QY 379 LeuGlnThrHisGlnLys 384
1018 CTGAGACACACCAAAAG 1035

RESULT 13
AC006501/c 147021 bp DNA linear PRI 26-FEB-1999
LOCUS Homo sapiens, clone hRPK.35_A_1, complete sequence.
DEFINITION AC006501
AC006501.5 GI:4309874
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
1 (bases 1 to 147021)
Birken, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens, clone hRPK.35_A_1
2 (bases 1 to 147021)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckely, R., Benn, J., Brown, A.,
Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,
Funks, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hagos, B., Heathford, A., Horton, L., Howland, J. C., Jones, C., Kam, L.,
Karatas, A., Lehoczeky, J., Liu, C., Locke, K., MacDonald, P.,
Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Nijhoff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Testaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.
Direct Submission
TITLE
JOURNAL
REFERENCE
AUTHORS
Submitted (03-FEB-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 147021)
Birken, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baker, J., Baldwin, J., Barna, N., Beckely, R., Benn, J., Brown, A.,
Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Depayre, E., Devon, K., Dewar, K.,
Donelan, L., Doyle, M., Ferreira, P., FitzHugh, W., Forrest, C.,
Funks, R., Gage, D., Galagan, J., Gardyna, S., Gilbert, D., Grant, G.,
Hagos, B., Heathford, A., Horton, L., Howland, J. C., Jones, C., Kam, L.,
Karatas, A., Lehoczeky, J., Liu, C., Locke, K., MacDonald, P.,
Marquis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, J.,
Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J.,
Naylor, J., Nijhoff, M., O'Connor, T., O'Donnell, P., Pavlin, B.,
Peterson, K., Pollara, V., Riley, R., Roberts, D., Roy, A., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Stone, C., Subramanian, A.,
Testaye, S., Torruella-Miller, I., Vassiliev, H., Vo, A., Wagner, A.,
Wheeler, J., Wu, X., Wyman, D., Ye, W. J. and Zody, M.

TITLE
JOURNAL
COMMENT

Wheeler, J., Wyman, D., Ye, W. J. and Zody, M.
Direct Submission
Submitted (16-FEB-1999) Whitehead Institute/MIT Center for Genome
Research, 330 Charles Street, Cambridge, MA 02141, USA
On Mar 1, 1999, this sequence version replaced gi:4309777.
All repeats were identified using RepeatMasker: Smit, A.F.A. &
Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>.

FEATURES

```

source
    1. 147021
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /clone="hrpx.35 A.1"
        /clone_1lb="RPC1-11 human BAC library"
        complement(382..683)
        /rpt_family="AluSx"
repeat_region
    1233..1446
        /rpt_family="L1MB3A"
repeat_region
    4058..4138
        /rpt_family="MER96"
        complement(4768..5069)
        /rpt_family="AluY"
        /note="Single-stranded coverage."
repeat_region
    5392..5693
        /rpt_family="AluSg"
        complement(5856..6314)
        /rpt_family="L1MC5"
repeat_region
    7499..7713
        /rpt_family="MIR"
        complement(9050..9111)
        /rpt_family="L2"
repeat_region
    9277..10125
        /rpt_family="AluYo/FRAM"
        complement(11036..11131)
        /rpt_family="MIR"
repeat_region
    12467..12556
        /rpt_family="MIR"
        complement(13039..13158)
        /rpt_family="MIR"
repeat_region
    13216..13368
        /rpt_family="MER5B"
        complement(13631..13682)
        /rpt_family="MIR"
repeat_region
    14013..14324
        /rpt_family="AluSg"
        complement(14881..15068)
        /rpt_family="L2"
repeat_region
    15777..15920
        /rpt_family="FLAM_A"
        complement(15977..16097)
        /rpt_family="MER5A"
repeat_region
    16111..16422
        /rpt_family="AluSx"
        complement(16435..16990)
        /rpt_family="L1MB2"
repeat_region
    16991..17293
        /rpt_family="AluSc"
        complement(17318..17483)
        /rpt_family="FAM"
repeat_region
    17491..17574
        /rpt_family="L1MB2"
        complement(17578..17632)
        /rpt_family="MER5B"
repeat_region
    18396..18544
        /rpt_family="MIR"
        complement(18601..18892)
        /rpt_family="AluSx"
        complement(21078..21102)
        /rpt_family="(CA)n"
repeat_region
    complement(21331..21814)
        /rpt_family="MLT1D"
        complement(24151..24326)
        /rpt_family="MLT1E1"
        complement(26673..26977)
        /rpt_family="AluSc"
        complement(27038..27349)
        /rpt_family="AluYb"
        complement(27614..27424)
        /rpt_family="MER3"
        complement(29218..28893)
        /rpt_family="MER7A"
        complement(31200..30893)
        /rpt_family="AluSg"
        complement(31363..31226)
        /rpt_family="MIR"
        complement(31608..31425)
        /rpt_family="MIR"
        complement(31715..31694)
        /rpt_family="AT rich"
        complement(32264..32559)
        /rpt_family="MER7A"
        complement(34056..34027)
        /rpt_family="(TTTG)n"
        complement(34057..34340)
        /rpt_family="AluSc"
        complement(34350..34643)
        /rpt_family="AluSx"
        complement(35919..36110)
        /rpt_family="MER34B"
        complement(36414..36113)
        /rpt_family="AluSg"
        complement(36626..36713)
        /rpt_family="MER34B"
        complement(37614..37317)
        /rpt_family="AluSx"
        complement(39201..39613)
        /rpt_family="L1MB3"
        complement(40549..40933)
        /rpt_family="MLT1G1"
        complement(41566..41531)
        /rpt_family="(TTTTG)n"
        complement(43494..43451)
        /rpt_family="(TC)n"
        complement(43767..44074)
        /rpt_family="AluSg"
        complement(44084..44352)
        /rpt_family="AluYb"
        complement(44359..44541)
        /rpt_family="L1MA6"
        complement(44872..44542)
        /rpt_family="AluSg"
        complement(44873..45183)
        /rpt_family="L1MA6"
        complement(45708..46376)
        /rpt_family="L1MB2"
        complement(46377..46677)
        /rpt_family="AluSx"
        complement(46678..46940)
        /rpt_family="L1MB2"
        complement(48115..48193)
        /rpt_family="MER5B"
        complement(48485..48681)
        /rpt_family="MIR"
        complement(49660..49373)
        /rpt_family="AluSx"
        complement(50181..50391)
        /rpt_family="MER58A"
        complement(52908..52831)
        /rpt_family="MIR"
        complement(57478..57191)
        /rpt_family="AluSc"
        complement(57518..57795)

```

```
repeat_region      /rpt_family="AluYb8"
                    5783..58280
repeat_region      /rpt_family="Alusg"
                    60124..60147
repeat_region      /rpt_family="(TTTG)n"
                    complement(60148..60393)
                    /rpt_family="Alusg"

Alignment Scores:
Pred. No.:         7,21e-55          Length:      147021
Score:            733.00             Matches:     176
Percent Similarity: 55.06%           Conservative: 36
Best Local Similarity: 45.71%         Mismatches:  98
Query Match:      35.81%             Indels:      75
DB:               Gaps:               11

US-10-077-111-13 (1-384) x AC006501 (1-147021)

QY      1 MetVallyleuilehsthrleuAlaAspHisGlyAspAspValaencCyvGAlaIaphe 20
        |||||
DB      106573 ATGATGAACCTGATTCACACATATAGCTGATCATGTGTACATGTCACATCTGCTGCTTC 106514

QY      21 SerPheSerleuAlaThrCySerSerleuAspLysThrIleArgleuTySerleuArg 40
        |||||
DB      106513 TCCCTTTCCCTCTGGCTACTGCTGCTTGGACAAAACAATTCGCTGTCTGTTACGT 106454

QY      41 AspPheThrGluIeuProHisSerProleuLysPheHisThrTyAlaValHisCyvGys 60
        |||||
DB      106453 GACTTATCTAGTACGACACATTCCTCATTTAGTTTCAATGCTATGCTGTCCATGCTGTC 106394

QY      61 CyvPheSerProSerGlyHisIleleuAlaSerCyvSerThrAspGlyThrValleu 80
        |||||
DB      106393 TGTTCCTCCCTTCAGACATATTTTGGCATTCGTTTCAACAAGATGTTACACTGCTCTA 106334

QY      81 TrpAsnThrGluIeuGlyIleMetleuAlaValMetGluIeuProSerGlySerProVal 100
        |||||
DB      106333 TGGAAATCTGAATAATGACACAGATCTGCGCATGTAGAAACGCTTATGCGACGCTCTG 106274

QY      101 ArgValCyvGlnPheSerProAspSerThrCyvleuAlaSerGlyAlaIaAspGlyThr 120
        |||||
DB      106273 AGGGTTGGCCAGTTTCCCAAGACTCCACGTTTGGCATTCAGGGGCGCATGTAGTAACT 106214

QY      121 ValValleuThrPheAlaGlnSerTyTyIleuTyArgCyvGlySerVallyAspGly 140
        |||||
DB      106213 GTGGTTTGTGGAATGACACAGTCATACAAATATATAGTAT-----GGA 106169

QY      141 SerleuAlaIaIaCyvAlaPheSerProAsnGlySerPheValThrGlySerSerCys 160
        |||||
DB      106168 TATTTAAATCTCTTCCTTTTAACTCTGTGACTAGGTGATCTTAAAAAAGTAAAGC 106109

QY      161 GlyAspLeuThrValTrpAspAspLysMetArgCyvleuHisSerGlyValaIaHisAsp 180
        |||||
DB      106108 AAAATATTTTGAATCTGACACAGCTCAAGTATGTAGAAAAAAGAAAAAATTTT--- 106052

QY      181 LeuGlyIleThrCyvAspPheSerSerGlnProValSerAspGlyIuGlnGlyLeu 200
        |||||
DB      106051 -----GATGACCCCTTAGAAGCAATAGTTTTCAGGTAGCAACAAGAA 106010

QY      201 GlnPhePheArgleuAlaSerCyvGlyGlnAspCyvGlnVallyIleThrIleValSer 220
        |||||
DB      106009 AAATATTTTAACTGGGCGCAGGTG-----TGTGCTCAACCGGTATATCCCAAGC 105959

QY      221 PheThrHisIleleuAlaArgArgThrGluHisGlnleuLysGlnPheThrGluAspTrp 240
        |||||
DB      105958 TTTCGGAGG-----CCAAGGACGCGGATCAAGAGTCAAGAGTCAAGAACCATCTCG 105905

QY      241 SerGluGluValAlaSerThrTrp-----LeuCyvAlaGlnAspLeuLysAsp-- 256
        |||||
DB      105904 -----CTAACATGCTGAACACCCCGCTCTCTACTAAAGATACAAAAATTAG 105860

QY      257 -----LeuValGlyIlePheLysMetAsnHisIleAsp-----GlyTyvGluLeu 271
        |||||
DB      105859 CCGGCGATGTGTGGACGCGCCTGTATGTCACAGTACTTGGAGGCTGAGCGACAGAGATG 105800
```

```
QY      272 LeuAsnLeuThrLysGluSerleuAla-----AspAspLeuLysIle-GluSerIe 288
        |||||
DB      105799 ACCTGAACCCAGGAGCGGAGCTTGCATGAGCCGAGATCACACCACTGATCTCCAGCCT 105740

QY      288 uGlyLeuArgSerLysValleuArgLysIleGluGluLeuArgThrLysVallySerIe 308
        |||||
DB      105739 GGGTGACAGA-----CCGAGCTCTCCGCTCAAA----- 105712

QY      308 uSerSerGlyIleProAspGluPheIleCyvProIleThrArgGluLeuMetLysAspR 328
        |||||
DB      105712 ----- 105712

QY      328 oValIleAlaSerAspGlyTyvSerTyGluLysGluAlaMetGluAsnTrpIleSerIy 348
        |||||
DB      105711 -----AAAAAAATTTTAACTGCGCTCTTAA 105680

QY      348 GlyTyvAspArgThrSerProMetThrAsnleuValleuProSerAlaValleuThrProAs 368
        |||||
DB      105679 AGGTCAAAAACCCAGAGAGTGTATATATGATTTATATAGTATGATTTGCTTAGCGCAA 105620

QY      368 natgThrleuLys 372
        |||||
DB      105619 TCTTAATAATTAG 105607

RESULT 14
AC096419/c 245795 bp DNA linear HTG 10-MAY-2003
LOCUS      Rattus norvegicus clone CH230-24D13, WORKING DRAFT SEQUENCE.
ACCESSION  AC096419
VERSION     AC096419.6 GI:30520867
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.

REFERENCE
1 (bases 1 to 245795)
AUTHORS   Muzny,D,Marle,,Metzker,M,lee,, Abramson,S,, Adams,C,, Alder,J,,
            Allen,C,, Allen,H,, Albrooks,S,, Amin,A,, Anguiano,D,,
            Anyalebechi,V,, Aoyagi,A,, Ayodeji,M,, Baca,E,, Baden,H,,
            Baldwin,D,, Bandaranaike,D,, Barber,M,, Barnstead,M,, Benahmed,F,,
            Biwalo,K,, Blair,J,, Blankenburg,K,, Blyth,P,, Brown,M,,
            Bryant,N,, Bunay,C,, Burch,P,, Burrell,K,, Calderon,E,,
            Cardenas,V,, Carter,K,, Cavazos,I,, Ceasar,H,, Center,A,,
            Chacko,J,, Chavez,D,, Chen,G,, Chen,R,, Chen,Y,, Chen,Z,, Chu,J,,
            Cleveland,C,, Cockrell,R,, Cox,C,, Coyle,M,, Crete,A,, d'Souza,L,,
            Davila,M,L,, Davis,C,, Davy-Carroll,L,, De Anda,C,, Dederich,D,,
            Delgado,O,, Denison,S,, Deramo,C,, Ding,Y,, Dinh,H,, Divya,K,,
            Draper,H,, Dugan-Rocha,S,, Dunn,A,, Durbin,K,, Duval,B,, Eaves,K,,
            Egan,A,, Escotto,M,, Eugene,C,, Evans,C,A,, Falls,T,, Fan,G,,
            Fernandez,S,, Finley,M,, Flagg,N,, Forbes,L,, Foster,M,, Foster,P,,
            Fraser,C,M,, Gabisi,A,, Ganta,R,, Garcia,A,, Garner,T,, Garza,M,,
            Gebregeorgis,E,, Geel,K,, Gill,R,, Grady,M,, Guerra,M,, Guevara,W,,
            Gunaratne,P,, Haaland,W,, Hamil,C,, Hamilton,K,, Hamilton,K,,
            Harvey,Y,, Havlak,P,, Hawes,A,, Henderson,N,, Hernandez,J,,
            Hernandez,R,, Hines,S,, Hladun,S,L,, Hodgson,A,, Hognes,M,,
            Hollins,B,, Howells,S,, Hulyk,S,, Hume,J,, Idlebird,D,, Jackson,A,,
            Jackson,L,, Jacob,L,, Jiang,H,, Johnson,B,, Johnson,R,, Jolyet,A,,
            Karpach,S,, Kelly,S,, Kelly,S,, Khan,Z,, King,L,, Kovar,C,,
            Kowsa,C,, Krat,C,L,, Lebow,H,, Levan,J,, Lewis,L,, Li,Z,, Liu,J,,
            Liu,J,, Liu,W,, Liu,Y,, London,P,, Longacre,S,, Lopez,J,,
            Lorensleu,L,, Louleaged,H,, Lozado,R,J,, Lu,X,, Ma,D,,
            Maheshwari,M,, Mahindaratne,M,, Mahmoud,M,, Malloy,K,, Mangum,A,,
            Mangum,B,, Mapa,P,, McNeill,T,Z,, Meenen,E,,
            Mawhinney,S,, McLeod,M,P,, McNeill,T,Z,, Meenen,E,,
            Milosavljevic,A,, Miner,G,, Minja,E,, Montemayor,J,, Moore,S,,
            Morgan,M,, Morris,K,, Morris,S,, Mundae,M,, Murphy,M,, Natf,L,,
            Nankervyls,C,, Neal,D,, Newton,N,, Nguyen,N,, Norris,S,,
            Nwackelmeb,O,, Okunou,G,, Olarnunnguen,A,, Pal,S,, Parks,K,,
            Pasternak,S,, Paul,H,, Perez,A,, Perez,L,, Pfannkoch,C,,
            Plopper,F,, Polndexter,A,, Popovic,D,, Primus,B,, Pu,L,-L,,
            Puzo,M,, Quiroz,J,, Rachlin,E,, Reeves,K,, Regier,M,A,, Reigh,R,,
```


Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
Rivers, C., Rodkey, T., Rojars, A., Rose, M., Rose, R., Ruiz, S. J.,
Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smajs, D.,
Sneed, A., Sodergren, E., Song, X.-Z., Sorrell, R., Sosa, J.,
Steinle, M., Strong, R., Sutton, A., Swalek, A., Taboc, P., Taylor, C.,
Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
Valas, R., Vera, V., Villaseca, D., Waldron, L., Walker, B., Wang, J.,
Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
Williams, G., Willson, R., Wleczek, R., Wooden, H., Worley, K.,
Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, Y., Yoon, V.,
Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
Niederhausern, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O.,
Weinstock, G. and Gibbs, R. A.

Direct Submission
Unpublished
2 (bases 1 to 245795)
Worley, K. C.

Direct Submission
Submitted (17-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 245795)
Rat Genome Sequencing Consortium.

Direct Submission
Submitted (10-MAY-2003) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

On May 10, 2003 this sequence version replaced gi:24819739.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GFNB
Center clone name: CH230-24D13
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 232948 bases at least Q40
Consensus quality: 234943 bases at least Q30
Consensus quality: 236332 bases at least Q20
Estimated insert size: 240759; sum-of-contigs estimation
Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_difft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contig. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 245795: contig of 245795 bp in length.
Location/Qualifiers
1. 245795
/organism="Rattus norvegicus"
/mol_type="genomic DNA"

FEATURES
source

/db xref="taxon:10116"
/clone="CH230-24D13"
1. 1225
misc_feature
/note="wgs_contig"
2242..3454
misc_feature
/note="wgs_contig"
243942..245795
misc_feature
/note="wgs_contig"

ORIGIN

Alignment Scores:
Pred. No.: 2,18e-54 Length: 245795
Score: 731.00 Matches: 164
Percent Similarity: 33.98% Conservative: 10
Best Local Similarity: 32.03% Mismatches: 20
Query Match: 35.71% Indels: 318
DB: 2 Gaps: 1

US-10-077-111-13 (1-384) x AC096419 (1-245795)

Qy 1 MetValValysLeuIleHsIThrLeuAlaAspHisGlyAspAspValaAsnCy6Cy6AlaPhe 20
Db 137442 ATGGTGAAGCTGATTCACACACTGGCGGATCAGGTGATGACGTCACTGCGGCTTC 137383

Qy 21 SerPheSerLeuLeuAlaThrCysSerLeuAspLysThrIleArgLeuTySerLeuArg 40
Db 137382 TCATCCACGCTCTCGGCGACCTGCTCTTGACAGACCATCCCGCTGATCTCCCTAAGT 137323

Qy 41 AspPheThrGluLeuProHisSerProLeuLysPheHisThrTyAlaValHisCysCys 60
Db 137322 GACTTTCGTAACCTGCGGCGACTCCCGTTGAAGTTTCACACGTAATGCGTCCACTGCTGC 137263

Qy 61 CysPheSerProSerGlyHisIleLeuAlaSerCysSerThrAspGlyThrThrValLeu 80
Db 137262 AGCTTTCGCGCGTGAAGACACGTTTGGCTGCTGCTGACAGACGCTACTACGTACTG 137203

Qy 81 TrpAnthrGluSerGlyGlyMetLeuAlaValMetGluGlnProSerGlySerProVal 100
Db 137202 TGGAGCCGACACGACGACACACCTTACCGTGTGAGACGACCGGAGGAGCCCTGTG 137143

Qy 101 ArgValCysGlnPheSerProAspSerThrCysLeuAlaSerGlyAlaAlaAspGlyThr 120
Db 137142 CGGTTTGTGCTTTTGGCGGACTCCACTTACGATCAGGAGGCTGCGATGGGTCT 137083

Qy 121 ValValLeuThrPheAlaGlnSerTyTyLeuTyTyr----- 132
Db 137082 GTTGTTTTGGGAACGACATTCATACAAATTATATGTAACGCGGCTTAAACCTGCC 137023

Qy 132 ----- 132
Db 137022 TGGCTGAGACGCCATTAAATTAATTTTCTTCACTGTGTAAATAATAGACATAGT 136963

Qy 132 ----- 132
Db 136962 TGAACCTATTAGGTAAATTAATTAAGTACTTGGAAATCGAATCCTTAGGTTAAGGT 136903

Qy 132 ----- 132
Db 136902 TGAATATCTCATGTCTCTCTGTTGTGTCTGCAGACGGGTGTGCCAGTGGCCAA 136843

Qy 132 ----- 132
Db 136842 GAGAGCTTGCCCTTTGTAGTCTCAGAGCTTAAGAAAGCCAGTCTTCAAGCATGTCCAG 136783

Qy 132 ----- 132
Db 136782 GGGTTAACCCTAGCTTAACAGCAGCAGCTTAAGAAAGTGTATATGTCTCCCTTTAGAA 136723

Qy 132 ----- 132
Db 136722 ACGATGCTTAGCTTAAGAGGGTTTTCACAGAGTGAAGTGTGATTTCCACCTGAGACAG 136663

Qy 132 ----- 132

Db 78089 GGAACGCGTTGCCAGGACGCCCTTTCGAACGCTTCCCGAGAGCTGCTGCTGC 78148
QY 132 ----- 132
Db 78149 ACATGTTTTTCCATAATGATTTGTTGAGATGAGCACAAGGCTTGTCTATTGTCTG 78208
QY 132 ----- 132
Db 78209 ATTGAATGCAGAAATCGAATGCGTCTTATTCTAATAGCATGTCTTTTGTAGAAA 78268
QY 132 ----- 132
Db 78269 TCATTCTGTAATTTAATGAGGCTTTGTTAGGCTTGCTGCTTTTGGCTTTTAAAG 78328
QY 132 ----- 132
Db 78329 CTTGGGGTTTAAATGCTCATATTGACTCGCATGAAAGTTAAGATACAAGAGTGT 78388
QY 132 ----- 132
Db 78389 AGGAGAGTCTCTGGATCTAGGCTCATAGGTATCAGCTGACGTGTTAAGACTCGCTG 78448
QY 132 ----- 132
Db 78449 TACTCTAAGGGGTGTACGGGTGTAGCATCTGATGCTTATTGTTAGGCAACGGTCACT 78508
QY 132 ----- 132
Db 78509 CTTGTTCTGTCAACGGCGGTTGTCAATTCTAAGGACTGTATAGAGGTCCAAAGCT 78568
QY 133 -----ArgCysGlySerValLysAspGlySerLeu 142
Db 78569 CTGAACCCCGGCTCTGCTCTGTTCTGTTTCAGGTGTGTGTCAAGGACAGCTCATTTG 78628
QY 143 AlaAlaCysAlaPheSerProAsnGlySerPhePheValThrGlySerSerCysGlyAsp 152
Db 78629 GTGGCTGTGCGTTTCTCTGACGAGGCTCTTGTCACTGGCTCTCTCGGAGGGGAC 78688
QY 163 LeuThrValTrpAspAspLysMetArgCysLeuHisSerGlyLysAlaHisAspLeuGly 182
Db 78689 TTGACGTGTGGGATGACAGATAGAGTGTCTGCACAGCAGAAAGCACAAGACTCGGG 78748
QY 183 IleThrCysCysAspPheSerSerGlnProValSer 194
Db 78749 ATCACCTGTGCAGCTTTCTCTCACAGCTCTCTCT 78784

Search completed: February 5, 2005, 13:45:30
Job time : 5876 secs